```
************************************************************
*                                                          *
*               M P S R C H                          (TM)  *
*                                                          *
************************************************************

          Release 3.1A John F. Collins, Biocomputing Research Unit.
                  Copyright (c) 1993-1998 University of Edinburgh, U.K.
                     Distribution rights by Oxford Molecular Ltd

MPsrch_PP      protein - protein database search, using Smith-Waterman algorithm

Run on:        Tue Apr 18 13:58:30 2000;   MasPar time 58.81 seconds
                                           339.279 Million cell updates/sec

Tabular output not generated.

Title:
Description:    >US-09-490-187-2
                (1-423) from US09490187.pep
Perfect Score:  3111
Sequence:       1 MALKVLLEQEKTFFILLVLL..........AVIHPATQSLQVRQRIGSL 423

Scoring table:  PAM 150
                Gap 11

Searched:       142080 seqs, 47172406 residues

Post-processing: Minimum Match 08
                 Listing first 45 summaries

Database:       pir62
                1:pir1 2:pir2 3:pir3 4:pir4

Statistics:     Mean 46.931; Variance 84.207; scale 0.557

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.


                                    SUMMARIES

Result                Query
  No.  Score  Match Length DB  ID        Description                          Pred. No.

   1   170    5.5    272    2  I48700    gene ox40 protein - m                2.81e-12
   2   149    4.8    271    2  S12783    OX40 antigen precurso                1.30e-08
   3   146    4.7    416    2  JN0006    nerve growth factor r                1.30e-08
   4   146    4.7    461    2  JC4302    tumor necrosis factor                4.18e-08
   5   142    4.6    454    1  GQMST1    tumor necrosis factor                1.95e-07
   6   142    4.6    461    2  I57826    tumor necrosis factor                1.95e-07
   7   143    4.6    455    1  GQHUT1    tumor necrosis factor                1.33e-07
   8   136    4.4    435    2  I54182    tumor necrosis factor                1.90e-06
   9   133    4.4    314    2  I37383    FAS soluble protein -                5.85e-06
  10   133    4.3    335    2  A40036    apoptosis-mediating s                5.85e-06
  11   130    4.2    326    1  GQVZML    T2 protein - myxoma v                1.77e-05
  12   130    4.2    461    2  GQRTT1    tumor necrosis factor                1.77e-05
  13   127    4.1    327    2  A46484    apoptosis-mediating m                5.31e-05
  14   128    4.1    349    2  D36858    gene G4R protein - va                3.69e-05
  15   126    4.1    459    2  I48854    gene murine tumour ne                7.63e-05
  16   126    4.1   1192    2  S69000    laminin gamma 2 chain                7.63e-05
  17   125    4.0    255    2  JT0752    lymphocyte activation                1.09e-04
  18   125    4.0    255    2  I38426    4-1BB - human                        1.09e-04
  19   125    4.0    474    2  B38634    tumor necrosis factor                1.09e-04
  20   122    3.9    324    2  JC23395   Fas antigen precursor                3.20e-04
  21   122    3.9    425    1  A26431    nerve growth factor r                3.20e-04
  22   121    3.9    427    1  GQHUN     nerve growth factor r                4.57e-04
  23   120    3.9    461    1  A35356    tumor necrosis factor                6.50e-04
  24   122    3.9   1111    2  B44018    laminin B2t chain - h                3.20e-04
  25   122    3.9   1193    2  A44018    laminin B2t chain pre                3.20e-04
  26   115    3.7   1680    2  A60771    furin (EC 3.4.21.75)                 3.71e-03
  27   111    3.6    277    2  S14458    B-cell activation pro                1.45e-02
  28   110    3.5   3075    2  MMMSA     laminin alpha-1 chain                2.02e-02
  29   108    3.5   3084    1  A49053    laminin alpha-1 chain                3.93e-02
  30   107    3.4    250    1  A45664    CD27 antigen precurso                5.46e-02
  31   105    3.4    256    2  B32293    T-cell antigen 4-1BB                 1.05e-01
  32   106    3.4    277    2  I37552    OX40 homolog - human                 7.58e-02
  33   105    3.4    349    2  T06680    hypothetical protein                 1.05e-01
  34   105    3.4    493    2  JC5486    membrane glycoprotein                1.05e-01
  35   105    3.4    703    2  A49994    thrombospondin 3 - mo                1.45e-01
  36   107    3.4   1548    2  S34583    sulfate transporter,                 7.58e-02
  37   106    3.4   2871    2  A55624    serine proteinase (EC                5.46e-02
  38   106    3.3    260    1  A45517    fibrillin-1 precurso                 7.58e-02
  39   104    3.3    289    2  A46515    CD27 antigen precurso                1.45e-01
  40   104    3.3    305    2  A46476    B cell-associated sur                1.45e-01
  41   103    3.3    596    1  A45664    CD40 - mouse                         2.00e-01
  42   104    3.3    956    1  A46016    variant-specific surf                2.00e-01
  43   103    3.3   2871    2  A55567    fibrillin - bovine                   2.00e-01
  44   103    3.3   3635    2  T10053    fibrillin I - human                  2.00e-01
  45   101    3.2   1607    1  MMMSB2    laminin gamma-1 chain                3.77e-01
```

```
                                    ALIGNMENTS

RESULT               1

ENTRY                I48700             #type complete
TITLE                gene ox40 protein - mouse
ALTERNATE_NAMES      OX40 antigen
ORGANISM             #formal_name Mus musculus #common_name house mouse
DATE                 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
                     18-Oct-1996
ACCESSIONS           I48700; I48334; S34377
REFERENCE            I48700
   #authors          Calderhead, D.M.; Buhlmann, J.E.; van den Eertwegh, A.J.;
                     Claasen, E.; Noelle, R.J.; Fell, H.P.
   #journal          J. Immunol. (1993) 151:5261-5271
   #title            Cloning of mouse Ox40: a T cell activation marker that may
                     mediate T-B cell interactions.
   #cross-references MUID:94044750
   #accession        I48700
     ##status        translated from GB/EMBL/DDBJ
     ##molecule_type mRNA
     ##residues      1-272 ##label RES
     ##cross-references EMBL:Z21674; NID:g312827; PID:g312828
REFERENCE            I48334
   #authors          Birkeland, M.L.; Copeland, N.G.; Gilbert, D.J.; Jenkins,
                     N.A.; Barclay, A.N.
   #journal          Eur. J. Immunol. (1995) 25:926-930
   #title            Gene structure and chromosomal localization of the mouse
                     homologue of rat OX40 protein.
   #cross-references MUID:95255413
   #accession        I48334
     ##status        translated from GB/EMBL/DDBJ
     ##molecule_type DNA
     ##residues      1-14,'G',16-272 ##label RE2
     ##cross-references EMBL:X85214; NID:g732818; PID:g732819
GENETICS
   #gene             ox40
   #introns          45/1; 86/1; 122/1; 144/2; 210/1; 250/1
   #length           272 #molecular-weight 30153 #checksum 841

SUMMARY          Best Local Similarity   5.5%;   DB 2;   Length 272;
                 Matches     48; Conservative   36; Mismatches   78; Indels   14; Gaps  11;

Qy   15 TLLVLLGYL:CKVTCESGDCRQQEFRDRSGNCVPCNOCGPGMELSKECGFGYGEDAQCVA 74
           :|| |:|  |  |:  | ||    ||   :| |  | || | :| ||||   ||   |
Db    9 TALILLA-L:LGVTARRLNCVKHTYPS-GHKC--CRECOPGHGMYSRCD--HTRDMLCHP 62

Qy   63 CETGFYNEAVNYDTCKQCTCQNHRSGSSELKQNCTPTQDTVC-RCRPGTQPROD-SGYKLG 120
           : || | |     |    |  :||| : :||   | |  |    |  ||    |  |:
Db   63 CETGFYNEAVNYDTCKQCTCQNHRSGSSELKQNCTPTQDTVC-RCRPGTQPROD-SGYKLG 120
```

```
Qy    75 CRLHRFKEDWGFQKCKPCLDCAVVNRFQ-KANCSAISDAICGDCLPGFYRKTKLVGFQ-D 132
Db   121 VDCVPCGPGHFSPGNNQACKPWTNCTLSGKQTRHPASDSLDAV-CED-RSLLATLL 174
Qy   133 MECVPC--GDPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALATVLLALL 186
```

RESULT   2

ENTRY   S12783   #type complete
TITLE   OX40 antigen precursor - rat
ALTERNATE_NAMES nerve growth factor receptor homolog
ORGANISM   #formal_name Rattus norvegicus #common_name Norway rat
DATE   30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
                20-Sep-1999
ACCESSIONS   S12783; S08036
REFERENCE   S12783; S08036
                #authors Mallett, S.; Fossum, S.; Barclay, A.N.
                #journal EMBO J. (1990) 9:1063-1068
                #title Characterization of the MRC OX40 antigen of activated CD4
                positive T lymphocytes - a molecule related to nerve growth
                factor receptor.
SUMMARY   #length 271 #molecular-weight 29895 #checksum 379
                #cross-references MUID:90214614
                #accession S12783
                ##molecule_type mRNA
                ##residues 1-271 ##label MAL
CLASSIFICATION #superfamily CD27 antigen; NGF receptor repeat homology
KEYWORDS   growth factor receptor; transmembrane protein
FEATURE
     1-19   #domain signal sequence #status predicted #label SIG\
    20-271   #product OX40 antigen #status predicted #label MAT\
   211-235   #domain transmembrane #status predicted #label TMM

Query Match   4.8%; Score 149; DB 2; Length 271;
Best Local Similarity 29.1%; Pred. No. 1.30e-08;
Matches 39; Conservative 27; Mismatches 57; Indels 11; Gaps 8;

```
Db     9 TAFLLLG-LSLGVTVKL-NCVKDTYPS-GHKC--CRECQPGHGMVSRCD--HTRDTVCHP 61
Qy    15 TLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVA 74
Db    62 CEPGFYNEAVNYDTCKQCTQCNHRGSELKQNCTPTEDTVC-QCRPGTQPRQDSSHKLGV 120
Qy    75 CRLHRFKEDWGFQKCKPCLDCAVVNRFQ-KANCSAISDAICGDCLPGFYRKTKLVGFQDM 133
Db   121 DCVPC--PPGHFSP 132
Qy   134 ECVPCGDPPPYEP 147
```

RESULT   3

ENTRY   JN0006   #type complete
TITLE   nerve growth factor receptor, low affinity precursor -
                chicken
ALTERNATE_NAMES NGF receptor
ORGANISM   #formal_name Gallus gallus #common_name chicken
DATE   10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change
ACCESSIONS   JN0006; A60504
REFERENCE   JN0006; A60504
                #authors Large, T.H.; Weskamp, G.; Helder, J.C.; Radeke, M.J.; Misko,
                T.P.; Shooter, E.M.; Reichardt, L.F.
                #journal Neuron (1989) 2:1123-1134
                #title Structure and developmental expression of the nerve growth
                factor receptor in the chicken central nervous system.
                #cross-references MUID:90016579
                #accession JN0006
                ##molecule_type mRNA
                ##residues 1-416 ##label LAR
REFERENCE   A60504
                #authors Heuer, J.G.; Fatemie-Nainie, S.; Wheeler, E.F.; Bothwell, M.
                #journal Dev. Biol. (1990) 137:287-304
                #title Structure and developmental expression of the chicken NGF
                receptor
                #cross-references MUID:90152140
                #accession A60504
                ##molecule_type mRNA
                ##residues 21-35,'Y',37-172,'K',174-275,'S',277-395,'R',397-416
                ##label HEU
COMMENT   preliminary; not compared with conceptual translation
COMMENT   This receptor is found on sensory and sympathetic neurons, on
                neuroblastoma cells, and on a variety of nonneuronal derivatives
                of the neural crest.
                The cysteine-rich region of the extracellular domain may form part
                or all of the NGF-binding site.
                This protein is thought to form a high-affinity receptor when it
                associates with the 140K trk proto-oncogene, which contains an
                intracellular tyrosine kinase domain.
CLASSIFICATION #superfamily nerve growth factor receptor; NGF receptor
                repeat homology
KEYWORDS   duplication; glycoprotein; heterodimer; monomer;
                phosphoprotein; receptor; transmembrane protein
FEATURE
     1-20   #domain signal sequence #status predicted #label SIG\
    21-416   #product nerve growth factor receptor #status predicted
                    #label MAT\
    21-239   #domain extracellular #status predicted #label EXT\
    24-57   #domain NGF receptor repeat homology #label NG1\
    59-100   #domain NGF receptor repeat homology #label NG2\
   101-139   #domain NGF receptor repeat homology #label NG3\
   141-181   #domain NGF receptor repeat homology #label NG4\
   189-237   #region serine/threonine-rich\
   240-261   #domain transmembrane #status predicted #label MEM\
   262-416   #domain intracellular #status predicted #label INT\
    52   #binding-site carbohydrate (Asn) (covalent) #status
                    predicted
SUMMARY   #length 416 #molecular-weight 44654 #checksum 3542

Query Match   4.7%; Score 146; DB 1; Length 416;
Best Local Similarity 30.0%; Pred. No. 4.18e-08;
Matches 54; Conservative 29; Mismatches 83; Indels 14; Gaps 13;

```
Db    90 PC-VESDDAVCRCAYGFYQDELSGSCKKECSICEVGFGIMPCR-D-SQDTVCEECPEGTF 146
Qy    24 SCKVTCESGDCR-QQE-FRDR-SGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRF 80
Db   147 SDEANFVDPCLPCTICEENEVMVK-ECTATSDAECRDLHPRWTHTHPPSLAGSDSPEPITR 205
Qy    81 KEDWGF-QKCKPCLDCAVVNRFQKANCSAISDAICGDCLPGFYRKT-KLVGFQDMECVPC 138
Db   206 DPFNTEGMATTLADIVTTVMGSSQPVVSRGTADNLIPVYCSILAAVVWGL-VAYIAF-KR 263
Qy   139 GD-PPPYEPHCASKVNLVKIASTASSPRDTA--LAAVICSALATVLLALLILCVIYCKR 195
```

RESULT   4

ENTRY   JC4302   #type complete
TITLE   tumor necrosis factor receptor p55 precursor - pig
ORGANISM   #formal_name Sus scrofa domestica #common_name domestic pig
DATE   29-Nov-1995 #sequence_revision 23-Jul-1999 #text_change
ACCESSIONS   JC4302; PC4093
REFERENCE   JC4302; PC4093
                #authors Suter, B.; Pauli, U.
                #journal Gene (1995) 163:263-266
                #title Cloning of the cDNA encoding the porcine p55 tumor necrosis
                factor receptor.
                #cross-references MUID:96011645
                #accession JC4302
                ##molecule_type mRNA
                ##residues 1-461 ##label SUT
                #cross-references GB:U19994; NID:g1141752; PIDN:AAC48499.1;
                PID:g1141753
                #accession PC4093

Column 1

```
      #molecule_type protein
      #residues 1-7 #label SU2
      #experimental_source kidney cell line 15
GENETICS
      #gene tnfr
CLASSIFICATION
      #superfamily tumor necrosis factor receptor type 1; NGF
      receptor repeat homology
KEYWORDS   glycoprotein; kidney; receptor; transmembrane protein; tumor
FEATURE
   1-29    #domain signal sequence #status predicted #label SIG\
   30-461  #product tumor necrosis factor receptor p55 #status
           predicted #label MAT\
   44-194  #domain extracellular cysteine rich #status predicted
   44-82   #domain NGF receptor repeat homology #label NG1\
   84-126  #domain NGF receptor repeat homology #label NGF\
   211-231 #domain transmembrane #status predicted #label TMM\
   361-447 #domain signal transduction #status predicted #label
           SIT\
   54,145,151 #binding_site carbohydrate (Asn) (covalent) #status
           predicted
SUMMARY    #length 461 #molecular-weight 50696 #checksum 8079
```

RESULT 5

```
Query Match        4.7%;  Score 146; DB 2; Length 461;
Best Local Similarity 25.5%;  Pred. No. 4.18e-08;
Matches 28;  Conservative 22;  Mismatches 55;  Indels 5;  Gaps 5;

Db   44 CPQGKYSHPQNRSICCTKCHKGTYLHNDC-LGPGLDTDCRECDNGTFPASENHLTQCLSC 102
           | ::| :|: .||:|| |:::| |   |:|
Qy   34 CRQQEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKE-DWGFQKCKPC 92

Db  103 SKCRSEMSQVEISPCTVDRDTVCG-CRKNQYRKYWSETLFQCLNCSLCPN 151
          |  ::  |:  :|  | ::|| |  || | |::  : |||
Qy   93 LDC-AVVNRFQKANCSATSDAICGDCLPGFYRKI-KLVGFQDMECVPCGD 140
```

```
ACCESSIONS A38634; B40254; S16677; S19021; I54532
REFERENCE A38634
   #authors Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice,
     G.C.; Wong, G.H.W.; Chen, E.Y.; Goeddel, D.V.
   #journal Proc. Natl. Acad. Sci. U.S.A. (1991) 88:2830-2834
   #title Cloning and expression of cDNAs for two distinct murine tumor
     necrosis factor receptors demonstrate one receptor is
     species specific.
   #cross-references MUID:91187885
   #accession A38634
   ##molecule_type mRNA
   ##residues 1-454 #label LEW
   ##cross-references GB:M60468; NID:g199825; PIDN:AAA39751.1; PID:g199826
ENTRY GQMST1        #type complete
TITLE   tumor necrosis factor receptor 1 precursor - mouse
ALTERNATE_NAMES  tumor necrosis factor receptor, 55K
ORGANISM  #formal_name Mus musculus #common_name house mouse
DATE  30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change
      22-Jun-1999
   #accession B40254
   ##molecule_type mRNA
   ##residues 1-454 #label GO2
REFERENCE A40254
   #authors Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan,
     C.I.; Copeland, N.G.; Jenkins, N.A.; Smith, C.A.
   #journal Mol. Cell. Biol. (1991) 11:3020-3026
   #title Molecular cloning and expression of the type 1 and type 2
     murine tumor necrosis factor receptors for tumor necrosis factor.
   #cross-references MUID:91246168
   #accession B40254
REFERENCE S16677
   #authors Barrett, K.; Taylor-Fishwick, D.A.; Cope, A.P.; Kissonerghis,
     A.M.; Gray, P.W.; Feldmann, M.; Foxwell, B.M.J.
   #journal Eur. J. Immunol. (1991) 21:1649-1656
   #title Cloning, expression and cross-linking analysis of the murine
     p55 tumor necrosis factor receptor.
```

Column 2

```
   #cross-references MUID:91285014
   #accession S16677
   ##molecule_type mRNA
   ##residues 1-454
   ##cross-references EMBL:X59238; NID:g53578; PIDN:CAA41922.1; PID:g53579
REFERENCE S19021
   #authors Rothe, J.G.; Brockhaus, M.; Gentz, R.; Lesslauer, W.
   #journal Immunogenetics (1991) 34:338-340
   #title Molecular cloning and expression of the mouse Tnf receptor
     type b.
   #cross-references MUID:92039815
   #accession S19021
   ##molecule_type mRNA
   ##residues 1-454 #label ROT
   ##cross-references EMBL:X57796; NID:g54848; PIDN:CAA40936.1; PID:g54849
REFERENCE I54532
   #authors Bebo, B.F.
   #journal Immunogenetics (1994) 39:450-451
   #title Nucleotide sequence of the TNF type I receptor from a mouse
     endothelioma cell line.
   #cross-references MUID:94245292
   #accession I54532
   ##status translated from GB/EMBL/DDBJ
COMMENT  This protein is one of two distantly related receptors for both
         TNF-alpha (cachectin) and TNF-beta (lymphotoxin).
CLASSIFICATION
   #superfamily tumor necrosis factor receptor type 1; NGF receptor
     repeat homology
KEYWORDS  glycoprotein; receptor; transmembrane protein
FEATURE
   1-29    #domain signal sequence #status predicted #label SIG\
   30-454  #product tumor necrosis factor receptor type 1 #status
           predicted #label MAT\
   30-212  #domain extracellular #status predicted #label EXT\
   44-82   #domain NGF receptor repeat homology #label NG1\
   84-126  #domain NGF receptor repeat homology #label NG2\
   127-167 #domain NGF receptor repeat homology #label NG3\
   168-204 #domain NGF receptor repeat homology #label NG4\
   213-235 #domain transmembrane #status predicted #label MEM\
   236-454 #domain intracellular #status predicted #label INT\
   54,151,202 #binding_site carbohydrate (Asn) (covalent) #status
           predicted
SUMMARY    #length 454 #molecular-weight 50129 #checksum 4839
```

RESULT 6

```
Query Match        4.6%;  Score 142; DB 1; Length 454;
Best Local Similarity 26.9%;  Pred. No. 1.95e-07;
Matches 29;  Conservative 16;  Mismatches 58;  Indels 5;  Gaps 5;

Db   44 CPQGKYVHSKNNSICCTKCHKGTYLVSDCPSP-GRDTVCRECEKGTFTASQNYLRQCLSC 102
Qy   34 CRQQEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKE-DWGFQKCKPC 92

Db  103 KTCRKEMSQVEISPCQADKDTVCG-CKENQFQRLSETHFQCVDCSPC 149
Qy   93 LDCA-VVNRFQKANCSATSDAICGDCLPG-FYRKTKLVGFQDMECVPC 138
```

```
ENTRY I57826        #type complete
TITLE   tumor necrosis factor receptor - mouse
ORGANISM  #formal_name Mus musculus #common_name house mouse
DATE  02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change
      23-Jul-1999
ACCESSIONS I57826; I57826
REFERENCE I57826
   #authors Rothe, J.G.; Bluethmann, H.; Gentz, R.; Lesslauer, W.;
     Steinmetz, M.
   #journal Mol. Immunol. (1993) 30:165-176
   #title Genomic organization and promoter function of the murine
     tumor necrosis factor receptor beta gene.
   #cross-references MUID:93156721
```

```
#accession      I57826
  ##status       preliminary; translated from GB/EMBL/DDBJ
  ##molecule_type DNA
  ##residues     1-454 #label RES
  ##cross-references GB:M76656; NID:g202100; PIDN:AAA40465.1; PID:g202102
GENETICS
  #gene          TNFR-2
  #introns       13/3; 65/1; 108/1; 158/1; 184/2; 210/1; 248/1; 257/3; 353/1
CLASSIFICATION  #superfamily tumor necrosis factor receptor type 1; NGF
                receptor repeat homology
KEYWORDS        cytokine receptor
FEATURE
  44-82          #domain NGF receptor repeat homology #label NGF
SUMMARY         #length 454 #molecular-weight 50030 #checksum 4267

Query Match    Best Local Similarity 26.9%;
Best Local Similarity 26.9%;  Score 142;  DB 2;  Length 454;
Matches 29;  Conservative 16;  Mismatches 58;  Indels 5;  Gaps 5;

Db     44 CPQGKYVHSKNNSICCTKCHKGTIVSDCPSP-GRDTVCRECEKGTFTASQNYLRQCLSC 102
Qy     34 CRQQEFRDSGNCVPCNQCCPGMELSKECGFGYGEDAQCVACRLHRFKEDWGFQK-CKPC 92
Db    103 KTCRKEMSQVEISPCQADKDTVCG-CKENQFQRYLSETHFQCVDCSPC 149
          ::::: :::: ::::: ::::::  ::::: ::::: : :::::
Qy     93 LDCA-VVNRFQKANCSATSDAICGDCLPG-FYRKTKLVGFQDMECVPC 138

RESULT    7
ENTRY          GQHUT1         #type complete
TITLE          tumor necrosis factor receptor 1 precursor - human
ALTERNATE_NAMES P55 tumor necrosis factor receptor; TNF receptor
CONTAINS       tumor necrosis factor alpha inhibitor; tumor necrosis factor
               binding protein 1 (TNF blocking factor)
ORGANISM       #formal_name Homo sapiens #common_name man
DATE           30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change
               22-Jun-1999
ACCESSIONS     A38208; A34899; A34900; A36555; C36555; A38281; S12057;
               JT0758; A60231; A38258; A60594; A35010; JC2404
REFERENCE      A38208
  #authors     Fuchs, P.; Strehl, S.; Dworzak, M.; Himmler, A.; Ambros, P.F.
  #journal     Genomics (1992) 13:219-224
  #title       Structure of the human TNF receptor 1 (p60) gene (TNFR1) and
               localization to chromosome 12p13.
  #cross-references MUID:92250049
  #accession   A38208
    ##molecule_type DNA
    ##residues  1-455 #label FUC
    ##cross-references GB:M75864; GB:M75865; GB:M75866; NID:g339748;
                PIDN:AAA61201.1; PID:g339750
REFERENCE      A34899
  #authors     Loetscher, H.; Pan, Y.C.E.; Lahm, H.W.; Gentz, R.; Brockhaus,
               M.; Tabuchi, H.; Lesslauer, W.
  #journal     Cell (1990) 61:351-359
  #title       Molecular cloning and expression of the human 55 kd tumor
               necrosis factor receptor.
  #cross-references MUID:90235284
  #accession   A34899
    ##molecule_type mRNA
    ##residues  1-455 #label LOE
    ##cross-references GB:M58286; GB:M33480; NID:g339753; PIDN:AAA36753.1;
                PID:g339754
  #experimental_source placenta
    ##note      part of this sequence, including the amino end of the
                mature protein, confirmed by protein sequencing
REFERENCE      A34900
  #authors     Schall, T.J.; Lewis, M.; Koller, K.J.; Lee, A.; Rice, G.C.;
               Wong, G.H.W.; Gatanaga, T.; Granger, G.A.; Lentz, R.; Raab,
               H.; Kohr, W.J.; Goeddel, D.V.
  #journal     Cell (1990) 61:361-370
  #title       Molecular cloning and expression of a receptor for human
               tumor necrosis factor.
  #cross-references MUID:90235285
```

```
#accession      A34900
  ##molecule_type mRNA
  ##residues     1-455 #label SCH
  ##cross-references GB:M33294; NID:g339744; PIDN:AAA03210.1; PID:g339745
REFERENCE      A36555
  #authors     Himmler, A.; Maurer-Fogy, I.; Kroenke, M.; Scheurich, P.;
               Pfizenmaier, K.; Lantz, M.; Olsson, I.; Hauptmann, R.;
               Stratowa, C.; Adolf, G.R.
  #journal     DNA Cell Biol. (1990) 9:705-715
  #title       Molecular cloning and expression of human and rat tumor
               necrosis factor receptor chain (p60) and its soluble
               derivative, tumor necrosis factor-binding protein.
  #cross-references MUID:91090841
  #accession   A36555
    ##molecule_type mRNA
    ##residues  1-455 #label HIM
    ##cross-references GB:M63121; NID:g339755; PIDN:AAA36754.1; PID:g339756
  #accession   C36555
    ##molecule_type protein
    ##residues  30-38;41-53;'X',55-79,'XX',82-94,'NK',;'XX',100-104;
                107-128;162-167,'X',169-201 #label HI2
    ##note      the purified protein, called tumor necrosis factor
                binding protein, is a soluble derivative of the
                receptor
REFERENCE      A38281
  #authors     Gray, P.W.; Barrett, K.; Chantry, D.; Turner, M.; Feldmann,
               M.
  #journal     Proc. Natl. Acad. Sci. U.S.A. (1990) 87:7380-7384
  #title       Cloning of human tumor necrosis factor (TNF) receptor cDNA
               and expression of recombinant soluble TNF-binding protein.
  #cross-references MUID:91017509
  #accession   A38281
    ##molecule_type mRNA
    ##residues  1-455 #label GRA
    ##cross-references GB:M77764
    ##note      the authors translated the codon TGG for residue 371 as
                Thr, AAG for residue 372 as Leu, and GAC for residue
                427 as Asn
REFERENCE      S12057
  #authors     Nophar, Y.; Kemper, O.; Brakebusch, C.; Engelmann, H.; Zwang,
               R.; Aderka, D.; Holtmann, H.; Wallach, D.
  #journal     EMBO J. (1990) 9:3269-3278
  #title       Soluble forms of tumor necrosis factor receptors (TNF-Rs).
               The cDNA for the type I TNF-R, cloned using amino acid
               sequence data of its soluble form, encodes both the cell
               surface and a soluble form of the receptor.
  #cross-references MUID:91006021
  #accession   S12057
    ##molecule_type mRNA
    ##residues  1-455 #label NOP
    ##cross-references EMBL:X55313; NID:g37223; PIDN:CAA39021.1; PID:g37224
    ##note      parts of soluble TNF binding protein 1, including its
                amino and carboxyl ends, were confirmed by protein
                sequencing
REFERENCE      JT0758
  #authors     Kemper, O.; Wallach, D.
  #journal     Gene (1993) 134:209-216
  #title       Cloning and partial characterization of the promoter for the
               human p55 tumor necrosis factor (TNF) receptor.
  #cross-references MUID:94085779
  #accession   JT0758
    ##molecule_type DNA
    ##residues  1-13 #label KEM
REFERENCE      A60231
  #authors     Seckinger, P.; Vey, E.; Turcatti, G.; Wingfield, P.; Dayer,
               J.M.
  #journal     Eur. J. Immunol. (1990) 20:1167-1174
  #title       Tumor necrosis factor inhibitor: purification, NH-2-terminal
               amino acid sequence and evidence for anti-inflammatory and
               immunomodulatory activities.
  #cross-references MUID:90292116
  #accession   A60231
    ##molecule_type protein
```

##residues   A38258   41-43,'X',45-53,'X',55-57 ##label SEC
REFERENCE
#authors   Gatanaga, T.; Hwang, C.; Kohr, W.; Cappuccini, F.; Lucci III,
            J.A.; Jeffes, E.W.B.; Lentz, R.; Tomich, J.; Yamamoto,
            R.S.; Granger, G.A.
#journal   Proc. Natl. Acad. Sci. U.S.A. (1990) 87:8781-8784
#title   Purification and characterization of an inhibitor (soluble
            tumor necrosis factor receptor) for tumor necrosis factor
            and lymphotoxin obtained from the serum ultrafiltrates of
            human cancer patients.
#cross-references MUID:91062364
#accession A38258
  ##molecule_type protein
  ##residues 41-60 ##label GAT
REFERENCE
#authors   Olsson, I.; Lantz, M.; Nilsson, E.; Peetre, C.; Thysell, H.;
            Grubb, A.; Adolf, G.
#journal   Eur. J. Haematol. (1989) 42:270-275
#title   Isolation and characterization of a tumor necrosis factor
            binding protein from urine.
#cross-references MUID:89171156
#accession A60594
  ##molecule_type protein
  ##residues 41-43,'X',45-53,'V',55-57,'XK',60 ##label OLS
  ##experimental_source renal failure patient urine
#cross-references MUID:90110215
#accession A35010
  ##molecule_type protein
  ##residues 41-45 ##label ENG
REFERENCE
#authors   Engelmann, H.; Novick, D.; Wallach, D.
#journal   J. Biol. Chem. (1990) 265:1531-1536
#title   Two tumor necrosis factor-binding proteins purified from
            human urine. Evidence for immunological cross-reactivity
            with cell surface tumor necrosis factor receptors.
#cross-references MUID:90128033
#accession JC2404
  ##molecule_type protein
  ##residues 41-53,'X',55-144,'X',146-150,'X',152-186,'X',188-201
  ##experimental_source urine
REFERENCE
#authors   Kajihara, J.; Asada, A.; Kirihara, S.; Kato, K.
#journal   Biosci. Biotechnol. Biochem. (1994) 58:2266-2268
#title   Amino acid sequence of natural tumor necrosis factor alpha
            inhibitor purified from human urine.
#cross-references MUID:95128033
#accession JC2404
  ##molecule_type protein
  ##residues 41-53,'X',55-144,'X',146-150,'X',152-186,'X',188-201
  ##label KAJ
  ##experimental_source normal urine
GENETICS
#gene   GDB:TNFR1
  ##cross-references GDB:125913; OMIM:191190
  ##map_position 12p13-12p13.2
CLASSIFICATION #superfamily tumor necrosis factor receptor type 1; NGF
            receptor repeat homology
KEYWORDS   duplication; glycoprotein; receptor; transmembrane protein
FEATURE
  1-21   #domain signal sequence #status predicted #label SIG\
  22-455   #product tumor necrosis factor receptor type 1 #status
             predicted #label MAT\
  30-211   #domain extracellular #status predicted #label EXT\
  41-201   #product TNF binding protein 1 (tumor necrosis factor
             alpha inhibitor) #status experimental #label TBP1\
  44-82   #domain NGF receptor repeat homology #label NG1\
  84-126   #domain NGF receptor repeat homology #label NG2\
  127-167   #domain NGF receptor repeat homology #label NG3\
  168-196   #domain NGF receptor repeat homology #label NG4\
  212-234   #domain transmembrane #status predicted #label MEM\
  235-455   #domain intracellular #status predicted #label INT\
  54,145,151 #binding_site carbohydrate (Asn) (covalent) #status
             predicted

COMMENT   This protein is one of two known receptors for both TNF-alpha
            (cachectin) and TNF-beta (lymphotoxin).

SUMMARY   #length 455 #molecular-weight 50494 #checksum 153

Query Match   4.6%;   Score 143; DB 1;   Length 455;
Best Local Similarity 25.2%; Pred. No. 1.33e-07;
Matches   32;   Conservative 30; Mismatches 57; Indels 8; Gaps 8;

Qy 11 LPL-VLLELVGIYPSGVIGLVPHLGDREKRDSV-CPQGKYIHPQNNSICCTKCHKGTYL 68
        :|: |||| :: ::|: :::|:||: |:   ::|: :|:|||:||:|: |:|:|:|||:
Db   1 MALKVLLEQ?KTFF-TLLV-LLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMEL 58

Db 69 YNDCP-GPGQDTDCRECESGSFFTASENHLRHCLSCSKCRKEMGQVEISSCTVDRDTVCG- 126
        | ||| :|:|: ::|:|:|::| | :|:| | |: :| ||: |||: ||: ||::|
Qy 59 SKECGFGYGEDAQCVACRLHRFKE-DWGFQKCKPCLDCA-VVNRFQKANCSATSDAICGD 116

Db 127 CRKNQYR 133
        :: ||
Qy 117 CLPGFYR 123

RESULT   8

ENTRY   I54182   #type complete
TITLE   tumor necrosis factor receptor 2-related protein - human
ORGANISM   #formal_name Homo sapiens #common_name man
DATE   24-May-1996 #sequence_revision 24-May-1996 #text_change
            29-Aug-1997
ACCESSIONS I54182
REFERENCE   I54182
#authors   Baens, M.; Chaffanet, M.; Cassiman, J.J.; Van den Berghe, H.;
            Marynen, P.
#journal   Genomics (1993) 16:214-218
#title   Construction and evaluation of a hncDNA library of human 12p
            transcribed sequences derived from a somatic cell hybrid.
#cross-references MUID:93252381
#accession I54182
  ##status preliminary; translated from GB/EMBL/DDBJ
  ##molecule_type mRNA
  ##residues 1-435 ##label RES
  ##cross-references GB:L04270; NID:g339761; PID:g339762
GENETICS
#gene   GDB:LTBR
  ##cross-references GDB:1230195; OMIM:600979
  ##map_position 12p13.3-12p13.1
SUMMARY   #length 435 #molecular-weight 46709 #checksum 63

Query Match   4.4%;   Score 136; DB 2;   Length 435;
Best Local Similarity 26.3%; Pred. No. 1.90e-06;
Matches   26;   Conservative 24; Mismatches 41; Indels 8; Gaps 8;

Db 37 ASENQTQTCRDQEKEYE-FQHRICCSRCPGPGTYSAKCS-RI-RDTVCATCAENSYNEHWN 93
        :|: |: |: :: :: :::| || |:| ||::: |: ::| |: :::|: :::|
Qy 28 TCESGDCR-LQQ-EFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKEDWG 85

Db 94 YLTICQLCRPCDPVMGLEEIAPCTSKRKTQC-RCQPGMF 131
        ::|:|::| :|:: | :: |:|: | ::| :||:
Qy 86 FQK-CKPCLDCAVVNRFQK-ANCSATSDAICGDCLPGFY 122

RESULT   9

ENTRY   I37383   #type complete
TITLE   FAS soluble protein - human
ORGANISM   #formal_name Homo sapiens #common_name man
DATE   02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
            02-Jul-1996
ACCESSIONS I37383
REFERENCE   I37383
#authors   Cascino, I.; Fiucci, G.; Papoff, G.; Ruberti, G.
#journal   J. Immunol. (1995) 154:2706-2713
#title   Three functional soluble forms of the human
            apoptosis-inducing Fas molecule are produced by alternative
            splicing.
#cross-references MUID:95181785
#accession I37383
  ##status preliminary; translated from GB/EMBL/DDBJ

RESULT 10

```
ENTRY        A40036        #type complete
TITLE        apoptosis-mediating surface antigen Fas precursor - human
ALTERNATE_NAMES surface antigen APO-1
ORGANISM     #formal_name Homo sapiens #common_name man
DATE         17-Jan-1992 #sequence_revision 17-Jan-1992 #text_change
             10-Sep-1997
ACCESSIONS   A40036; S24543; A38142
REFERENCE    A40036
   #authors  Itoh, N.; Yonehara, S.; Ishii, A.; Yonehara, M.; Mizushima,
             S.I.; Sameshima, M.; Hase, A.; Seto, Y.; Nagata, S.
   #journal  Cell (1991) 66:233-243
   #title    The polypeptide encoded by the cDNA for human cell surface
             antigen Fas can mediate apoptosis.
   #cross-references MUID:91309137
   #accession A40036
             ##molecule_type mRNA
             ##residues 1-335 ##label KRA
             ##cross-references EMBL:X63717; NID:g28741; PID:g28742
REFERENCE    A38142
   #authors  Oehm, A.; Behrmann, I.; Falk, W.; Pawlita, M.; Maier, G.;
             Klas, C.; Li-Weber, M.; Richards, S.; Dhein, J.; Trauth,
             B.C.; Ponstingl, H.; Krammer, P.H.
   #journal  J. Biol. Chem. (1992) 267:10709-10715
   #title    Purification and molecular cloning of the APO-1 cell surface
             antigen, a member of the tumor necrosis factor/nerve growth
             factor receptor superfamily. Sequence identity with the Fas
             antigen
   #accession A38142
             ##molecule_type mRNA
             ##residues 1-335 ##label OEH
             ##cross-references EMBL:M67454; NID:g182409; PID:g182410
GENETICS
   #gene     GDB:AP1
             #map_position 10q24.1-10q24.1
             #cross-references GDB:132671; OMIM:134637
CLASSIFICATION #superfamily NGF receptor repeat homology
KEYWORDS     apoptosis; surface antigen; transmembrane protein
FEATURE
     1-16    #domain signal sequence #status predicted #label SIG\
    85-128   #domain NGF receptor repeat homology #label NG4\
   174-190   #domain transmembrane #status predicted #label TMM
SUMMARY      #molecule_type mRNA
             #residues 1-314 ##label RES
             ##cross-references EMBL:Z47993; NID:g728578; PID:g695539
             #length 314 #molecular-weight 35386 #checksum 655

Query Match      4.3%; Score 133; DB 2; Length 314;
Best Local Similarity  26.9%; Pred. No. 5.85e-06;
Matches  21; Conservative  21; Mismatches 30; Indels  6; Gaps  6;

Db   59 CHRKPCPPGERKARDCTVN-GDEPDCVPCQEGKEYTDKAHFSSKCRRCRLCDEGHGLEVEI 117
            ::: ||| : ::|:|:||||::|||::|| | ::|
Qy   49 CNQ-CGPGMELSKECGFGYGEDAQCVACRLHR-FKEDWGFQ-RCKPCLDCAVVNRFQ-KA 104

Db  118 NCTRTQNTKC-RCKPNFF 134
            |:|| |:| ||
Qy  105 NCSATSDAICGDCLPGFY 122
```

RESULT 11

```
ENTRY        GQVZML        #type complete
TITLE        T2 protein - myxoma virus (strain Lausanne)
ORGANISM     #formal_name myxoma virus
DATE         31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change
             18-Jun-1999
ACCESSIONS   A40566
REFERENCE    A40566
   #authors  Upton, C.; Macen, J.L.; Schreiber, M.; Mcfadden, G.
   #journal  Virology (1991) 184:370-382
   #title    Myxoma virus expresses a secreted protein with homology to
             the tumor necrosis factor receptor gene family that
             contributes to viral virulence.
   #cross-references MUID:91335768
   #accession A40566
             ##molecule_type DNA
             ##residues 1-326 ##label UPT
             ##cross-references GB:M95181; GB:M37976; NID:g333309; PIDN:AAA46632.1;
             PID:g333310
CLASSIFICATION #superfamily myxoma virus T2 protein; NGF receptor repeat
             homology
KEYWORDS     glycoprotein
FEATURE
    64-105   #domain NGF receptor repeat homology #label NG2\
   106-147   #domain NGF receptor repeat homology #label NG3\
   66,181,205,238  #binding-site carbohydrate (Asn) (covalent) #status
             predicted
SUMMARY      #length 326 #molecular-weight 35208 #checksum 9255

Query Match      4.3%; Score 130; DB 1; Length 326;
Best Local Similarity  29.5%; Pred. No. 1.77e-05;
Matches  28; Conservative  16; Mismatches 42; Indels  9; Gaps  8;

Db   24 DRGKCRGNDY-EKDGLC--CTSCPPGSYASRLCG-P-GSDTVCGSPCKNETFTASTNHAPA 78
Qy   30 ESGDCRQEFRDRSGNCVPCNCQGPGMELSKECGFGYGEDAQCVACRLHRFKEDWGF-QK 88

Db   79 CVSCRGRCTG-HLSESOSCDKTRDRVC-DCSAGNY 111
Qy   89 CKPCLD-CAVVNRFQKANCSATSDAICGDCLPGFY 122
```

RESULT 12

```
ENTRY        GQRTT1        #type complete
TITLE        tumor necrosis factor receptor 1 precursor - rat
CONTAINS     tumor necrosis factor receptor binding protein 1 (TNF blocking factor)
ORGANISM     #formal_name Rattus norvegicus #common_name Norway rat
DATE         30-Jun-1992 #sequence_revision 07-Oct-1994 #text_change
             22-Jun-1999
ACCESSIONS   B36555
REFERENCE    B36555
   #authors  Himmler, A.; Maurer-Fogy, I.; Kroenke, M.; Scheurich, P.;
             Pfizenmaier, K.; Lantz, M.; Olsson, I.; Hauptmann, R.;
             Stratowa, C.; Adolf, G.R.
   #journal  DNA Cell Biol. (1990) 9:705-715
   #title    Molecular cloning and expression of human and rat tumor
             necrosis factor receptor chain (p60) and its soluble
             derivative, tumor necrosis factor-binding protein.
SUMMARY      #length 335 #molecular-weight 37732 #checksum 4899

Query Match      4.3%; Score 133; DB 2; Length 335;
Best Local Similarity  26.9%; Pred. No. 5.85e-06;
Matches  21; Conservative  21; Mismatches 30; Indels  6; Gaps  6;

Db   59 CHRKPCPPGERKARDCTVN-GDEPDCVPCQEGKEYTDKAHFSSKCRRCRLCDEGHGLEVEI 117
Qy   49 CNQ-CGPGMELSKECGFGYGEDAQCVACRLHR-FKEDWGFQ-RCKPCLDCAVVNRFQ-KA 104

Db  118 NCTRTQNTKC-RCKPNFF 134
Qy  105 NCSATSDAICGDCLPGFY 122
```

```
#cross-references MUID:91090841
  ##accession B36555
  ##molecule_type mRNA
  ##residues 1-461 ##label HIM
#cross-references GB:M63122; NID:g207361; PIDN:AAA42256.1; PID:g207362
COMMENT    This protein is one of two known receptors for both TNF-alpha
           (cachectin) and TNF-beta (lymphotoxin).
CLASSIFICATION #superfamily tumor necrosis factor receptor type 1; NGF
               receptor repeat homology
KEYWORDS   duplication; glycoprotein; receptor; transmembrane protein
FEATURE
   1-29           #domain signal sequence #status predicted #label SIG\
   30-461         #product tumor necrosis factor receptor type 1 #status
                  predicted #label MAT\
   30-211         #domain extracellular #status predicted #label EXT\
   30-201         #product tumor necrosis factor binding protein #status
                  predicted #label TBP\
   44-82          #domain NGF receptor repeat homology #label NG1\
   84-126         #domain NGF receptor repeat homology #label NG2\
   127-167        #domain NGF receptor repeat homology #label NG3\
   168-204        #domain NGF receptor repeat homology #label NG4\
   212-234        #domain transmembrane #status predicted #label MEM\
   235-461        #domain intracellular #status predicted #label INT\
   54,151,201     #binding_site carbohydrate (Asn) (covalent) #status
                  predicted
SUMMARY    #length 461 #molecular-weight 50969 #checksum 1617

Query Match        4.2%;  Score 130;  DB 1;  Length 461;
Best Local Similarity 25.9%;  Pred. No. 1.77e-05;
Matches   28;  Conservative  16;  Mismatches  59;  Indels   5;  Gaps   5;

Db   44 CPQGKYAHPKNNSICCTKCHKGTYLVSDCPSP-GQETVCEVCDKGTFTASQNHVRQCLSC 102
           | |  : |  ::|     ::   |    :| |   |   | : | : :|
Qy   34 CRQQEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKEDWGF-QKCKPC 92

Db  103 KTCRKEMFQVEISPCKADMDTVCG-CKKNQFQRYLSETHFQCVDCSPC 149
           :.: |: :|  :  ||::||   |  | |  | |    || | |
Qy   93 LDCA-VVNRFQKANCSATSDAICGDCLPG-FYRKTKLVGFQDMECVPC 138
```

```
RESULT  13
ENTRY      A46484        #type complete
TITLE      apoptosis-mediating membrane-associated polypeptide Fas -
           mouse
ORGANISM   #formal_name Mus musculus #common_name house mouse
DATE       18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change
           16-Jul-1999
ACCESSIONS A46484; A47254
REFERENCE  A46484
           #authors Watanabe-Fukunaga, R.; Brannan, C.I.; Itoh, N.; Yonehara, S.;
                    Copeland, N.G.; Jenkins, N.A.; Nagata, S.
           #journal J. Immunol. (1992) 148:1274-1279
           #title The cDNA structure, expression, and chromosomal assignment of
                  the mouse Fas antigen.
  #cross-references MUID:92148151
    ##accession A46484
    ##status preliminary
    ##molecule_type mRNA
    ##residues 1-327 ##label WAT
  #cross-references GB:M83649; NID:g193225; PID:g193226
    ##experimental_source BAM3 macrophage cell line
    ##note sequence extracted from NCBI backbone (NCBIN:81544,
           NCBIP:81545)
REFERENCE  A47254
           #authors Adachi, M.; Watanabe-Fukunaga, R.; Nagata, S.
           #journal Proc. Natl. Acad. Sci. U.S.A. (1993) 90:1756-1760
           #title Aberrant transcription caused by the insertion of an early
                  transposable element in an intron of the Fas antigen gene
                  of lpr mice.
  #cross-references MUID:93189576
    ##accession A47254
    ##status preliminary
    ##molecule_type nucleic acid
```

```
  ##residues 1-96 ##label ADA
  ##cross-references GB:S56490; NID:g2985505; PID:g2985506
  ##experimental_source MRL lpr/lpr
  ##note sequence extracted from NCBI backbone (NCBIN:126850,
         NCBIN:126853, NCBIN:126863, NCBIP:126864)
CLASSIFICATION #superfamily NGF receptor repeat homology
               transmembrane protein
KEYWORDS
FEATURE
   44-79          #domain NGF receptor repeat homology #label NGF\
   81-124         #domain NGF receptor repeat homology #label NG4\
SUMMARY    #length 327 #molecular-weight 37417 #checksum 8479

Query Match        4.1%;  Score 127;  DB 2;  Length 327;
Best Local Similarity 26.0%;  Pred. No. 5.31e-05;
Matches   20;  Conservative  18;  Mismatches  34;  Indels   5;  Gaps   5;

Db   56 CQPCQPGKKIVEDCKRMN-GGTPTCAPCTEGKEYMDKNHYADKCRRCTLCDEEHGLEVETN 114
           | ||       :|    :| |  |  |  :|  |  :   |  :||   || ||   :| |
Qy   49 CNQCGPGME.SKECGFGYGEDAQCVACRLHR-FKEDWGF-QKCKPCLDCAVVNRFQ-KAN 105

Db  115 CTLTQNTKC-KCKPGFY 130
           |  :  | | |  | ||
Qy  106 CSATSDAICGDCLPGFY 122
```

```
RESULT  14
ENTRY      D36858        #type complete
TITLE      gene G4R protein - variola virus
ALTERNATE_NAMES B23R protein (COP)
ORGANISM   #formal_name variola virus
DATE       30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
           26-Aug-1999
ACCESSIONS D36858; S46888; S32385; S35987
REFERENCE  A36859
           #authors Blinov, V.M.
           #submission submitted to GenBank, November 1992
           #description no_shown.
  #accession D36858
    ##status preliminary
    ##molecule_type DNA
    ##residues 1-349 ##label BLI
    ##cross-references GB:X69198; NID:g456758; PID:g457087
    ##experimental_source strain India-1967, ssp. major, isolate Ind3
REFERENCE  S46868
           #authors Kolykhalov, A.A.; Blinov, V.M.; Gytorov, V.V.; Pozdnyakov,
                    S.G.; Chizhikov, V.E.; Frolov, I.V.; Totmenin, A.V.;
                    Shchelkunov, S.N.; Sandakhchiev, L.S.
           #submission submitted to the EMBL Data Library, April 1992
           #description Nucleotide sequence analysis of the region of variola virus
                        XhoI F O H P Q genome fragment.
  #accession S46888
    ##status preliminary
    ##molecule_type DNA
    ##residues 1-349 ##label KOL
    ##cross-references EMBL:X67117; NID:g516428; PID:g516449
    ##experimental_source strain India-1967, isolate Ind3
REFERENCE  S32385
           #authors Shchelkunov, S.N.; Blinov, V.M.; Sandakhchiev, L.S.
           #journal FEBS Lett. (1993) 319:80-83
           #title Genes of variola and vaccinia viruses necessary to overcome
                  the host protective mechanisms.
  #cross-references MUID:93202281
    ##accession S32385
    ##status preliminary
    ##molecule_type DNA
    ##residues 31-168 ##label SHC
    ##cross-references EMBL:X69198
    ##experimental_source strain India-1967, ssp. major
GENETICS
  #gene G4R
CLASSIFICATION #superfamily myxoma virus T2 protein; NGF receptor repeat
               homology
FEATURE
   32-66          #domain NGF receptor repeat homology #label NGF\
```

```
            68-109      #domain NGF receptor repeat homology #label NG2\
           110-151      #domain NGF receptor repeat homology #label NG3
SUMMARY                 #length 349 #molecular-weight 38189 #checksum 2016

Query Match             4.18;    Score 128;  DB 2;  Length 349;
Best Local Similarity  28.7%;    Pred. No. 3.69e-05;
Matches    27;  Conservative  21;  Mismatches  37;  Indels   9;  Gaps   8;

Db  29  NGCKDTEYK-RHNIC--CLSCPPGTYASRLCD-S-KTNTQCTPCGSGTFTSRNNHLPAC  83
        :|:||||||| :||::||::||| :||    |      :||:| |  :||  |||  :||
Qy  31  SGDCRQQEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKE-DWGFQKC  89

Db  84  LSCNGRCNS-NQVETRSCNTTHNRIC-ECSPGYY  115
        ::|   |:  ||: :|: :|| ::| | ||:||:
Qy  90  KPCLD-CAVVNRFQKANCSATSDAICGDCLPGFY  122

RESULT  15
ENTRY           148854      #type fragment
TITLE           gene murine tumour necrosis factor receptor 2 protein - mouse
                (fragment)
ORGANISM        #formal_name Mus musculus #common_name house mouse
DATE            02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
                23-Jul-1999
ACCESSIONS      I48854
REFERENCE       I48854
                #authors Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.
                #journal Mamm. Genome (1994) 5:726-727
                #title Allelic variation of the type 2 tumor necrosis factor
                receptor gene.
                #cross-references MUID:95178848
                #accession I48854
                ##status preliminary; translated from GB/EMBL/DDBJ
                ##molecule_type mRNA
                ##residues 1-459 #label RES
                ##cross-references EMBL:X76401; NID:g433830; PIDN:CAA53981.1;
                PID:g433831

CLASSIFICATION  #superfamily tumor necrosis factor receptor type 2; NGF
                receptor repeat homology

FEATURE
           151-188      #domain NGF receptor repeat homology #label NGF
SUMMARY                 #length 459 #checksum 3156

Query Match             4.18;    Score 126;  DB 2;  Length 459;
Best Local Similarity  25.0%;    Pred. No. 7.63e-05;
Matches    24;  Conservative  20;  Mismatches  45;  Indels   7;  Gaps   7;

Db  20  EPGYECQISQEYDRKAQ-MCCAKCPPGQYVKHFCN-K-TSDTVCADCEASMTQVWNQF  76
        | | :|: ||: |:: ||  || |   |  :||| |   ||| :|| | | : | |||
Qy  30  ESG-DCR-QQEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKEDWG-F  86

Db  77  RTCLSCSSSCSTDQVETRACTKQQNRVCA-CEAGRY  111
        :||  | | :| ||:|| ||::|:|| || |   :|
Qy  87  QKCKPCLDCAVVNRFQKANCSATSDAICGDCLPGFY  122

Search completed: Tue Apr 18 13:59:33 2000
Job time : 63 secs.
```

```
*************************************************
                  M P S R C H  (TM)
*************************************************

      Release 3.1A John F. Collins, Biocomputing Research Unit.
      Copyright (c) 1993-1998 University of Edinburgh, U.K.
      Distribution rights by Oxford Molecular Ltd

MPsrch_pp    protein - protein database search, using Smith-Waterman algorithm

Run on:          Tue Apr 18 14:02:47 2000;   MasPar time 7.76 Seconds
                                              705.843 Million cell updates/sec

Tabular output not generated.

Title:
Description:     >US-09-490-187-2
                 (1-423) from US09490187.pep
Perfect Score:   3111
Sequence:        1 MALKVLLEQEKTEFTLIVLL..........AVIHPATQTSLQVRQRLGSL 423

Scoring table:   PAM 150
                 Gap 11

Searched:        131253 seqs, 12956647 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        a-issued
                 1:5A_COMB 2:5B_COMB 3:PCT9_COMB 4:backfiles1

Statistics:      Mean 32.744;  Variance 152.615;  scale 0.215

                 Pred. No. is the number of results predicted by chance to have a
                 score greater than or equal to the score of the result being printed,
                 and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	187	6.0	186	1	US-08-089-	Sequence 6, Applicatio	3.49e-06
2	178	5.7	206	1	US-08-494-	Sequence 7, Applicati	1.81e-05
3	178	5.7	206	1	US-08-097-	Sequence 7, Applicati	1.81e-05
4	178	5.7	438	1	US-08-494-	Sequence 11, Applicati	1.81e-05
5	178	5.7	438	1	US-08-097-	Sequence 11, Applicati	1.81e-05
6	151	4.9	158	1	US-08-494-	Sequence 54, Applicati	2.27e-03
7	151	4.9	158	1	US-08-465-	Sequence 54, Applicati	2.27e-03
8	148	4.8	139	2	US-08-219-	Sequence 8, Applicati	3.84e-03
9	148	4.8	355	1	US-08-292-	Sequence 6, Applicati	3.84e-03
10	147	4.7	154	2	US-08-232-	Sequence 10, Applicati	4.58e-03
11	143	4.6	167	1	US-08-465-	Sequence 57, Applicati	9.20e-03
12	143	4.6	167	2	US-08-050-	Sequence 2, Applicati	9.20e-03
13	143	4.6	167	1	US-08-465-	Sequence 57, Applicati	9.20e-03
14	143	4.6	167	2	US-08-050-	Sequence 2, Applicati	9.20e-03
15	143	4.6	455	2	US-08-465-	Sequence 25, Applicati	9.20e-03
16	143	4.6	455	2	US-08-126-	Sequence 25, Applicati	9.20e-03
17	143	4.6	455	2	US-08-837-	Sequence 2, Applicati	9.20e-03
18	143	4.6	455	1	US-08-321-	Sequence 2, Applicati	9.20e-03
19	143	4.6	455	1	US-08-050-	Sequence 2, Applicati	9.20e-03
20	138	4.4	124	1	US-08-050-	Sequence 4, Applicati	2.19e-02
21	138	4.4	124	1	US-08-465-	Sequence 4, Applicati	2.19e-02
22	138	4.4	153	2	US-08-219-	Sequence 4, Applicati	2.19e-02
23	136	4.4	197	2	US-08-505-	Sequence 1, Applicatio	3.09e-02
24	133	4.3	119	2	US-08-219-	Sequence 3, Applicatio	5.17e-02
25	134	4.3	199	2	US-08-465-	Sequence 48, Applicati	4.35e-02
26	134	4.3	199	1	US-08-050-	Sequence 48, Applicati	4.35e-02
27	133	4.3	314	1	PCT-US95-1	Sequence 5, Applicatio	5.17e-02
28	133	4.3	314	1	US-08-444-	Sequence 19, Applicati	5.17e-02
29	133	4.3	314	3	US-08-152-	Sequence 19, Applicati	5.17e-02
30	133	4.3	335	3	PCT-US95-1	Sequence 2, Applicatio	5.17e-02
31	133	4.3	335	2	US-08-219-	Sequence 2, Applicatio	5.17e-02
32	133	4.3	335	2	US-08-050-	Sequence 1, Applicatio	5.17e-02
33	130	4.3	326	3	US-08-409-	Sequence 1, Applicatio	8.62e-02
34	130	4.2	326	1	PCT-US91-0	Sequence 8, Applicati	8.62e-02
35	128	4.1	283	3	US-08-292-	Sequence 4, Applicati	1.21e-01
36	128	4.1	153	1	US-08-050-	Sequence 2, Applicati	1.21e-01
37	125	4.0	153	2	US-08-465-	Sequence 52, Applicati	2.01e-01
38	125	4.0	255	2	US-08-816-	Sequence 9, Applicati	2.01e-01
39	125	4.0	255	3	PCT-US96-0	Sequence 8, Applicati	2.01e-01
40	125	4.0	255	1	US-08-236-	Sequence 8, Applicati	2.01e-01
41	125	4.0	474	4	5395760-4	Patent NO. 5395760.	2.01e-01
42	125	4.0	474	2	US-08-650-	Sequence 4, Applicati	2.01e-01
43	122	3.9	1111	1	US-08-317-	Sequence 15, Applicati	3.32e-01
44	122	3.9	1171	1	US-08-445-	Sequence 1, Applicati	3.32e-01
45	122	3.9	1193	1	US-08-317-	Sequence 13, Applicati	3.32e-01

ALIGNMENTS

```
RESULT    1
ID   US-08-089-458B-6    STANDARD;    PRT;   186 AA.
XX
AC   XXXXXX
XX
DT
XX
DE   Sequence 6, Application US/08089458B
XX
CC   Patent NO. 5359039
CC   Sequence 6, Application US/08089458B
CC   GENERAL INFORMATION:
CC   APPLICANT:  Smith, Craig
CC   APPLICANT:  Goodwin, Raymond
CC   TITLE OF INVENTION:  Isolated Poxvirus A53R-Equivalent Tumor
CC   TITLE OF INVENTION:  Necrosis Factor Antagonists
CC   NUMBER OF SEQUENCES:  7
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE:  Patricia Anne Perkins, Immunex Corporation
CC   STREET:  51 University Street
CC   CITY:  Seattle
CC   STATE:  Washington
CC   COUNTRY:  U.S.A.
CC   ZIP:  98101
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE:  Floppy disk
CC   COMPUTER:  Apple Macintosh
CC   OPERATING SYSTEM:  Apple Operating System 7.1
CC   SOFTWARE:  Microsoft Word, Version #5.1a
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER:  US/08/089,458B
CC   FILING DATE:  07/09/93
CC   CLASSIFICATION:  514
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME:  Perkins, Patricia A.
CC   REGISTRATION NUMBER:  34,693
CC   REFERENCE/DOCKET NUMBER:  2608
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE:  (206) 587-0430
CC   TELEFAX:  (206) 233-0644
CC   TELEX:  756822
CC   INFORMATION FOR SEQ ID NO:  6:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH:  186 amino acids
CC   TYPE:  amino acid
CC   TOPOLOGY:  linear
CC   MOLECULE TYPE:  protein
```

```
SQ  SEQUENCE 186 AA; 20372 MW; 170907 CN;

Query Match           6.0%;  Score 187;  DB 1;  Length 186;
Best Local Similarity 30.0%;  Pred. No. 3.49e-06;
Matches   33;  Conservative  28;  Mismatches  40;  Indels   9;  Gaps   9;

Qy   34 NGSCDDGEYLDKTHN-QCCNRCPPG-EFAKIRCS-G-SDNTKCERCPPHYTIVPNYSNG  89
Db   90 CHQCRKCPT-GSFDKVKCTGTQNSKCS-CLPGWFCCATDSSKTEDCRDCIP 137
Qy   31 SGDCRQQEFPRDRSGNCVPCNQCGPGMELSKE-CGFGYGEDAQCVACRLHRFKEDWGFQK-  88
Db   89 CKPCLDCAVVNRFQKANCSATSDAICGDCLPGFYRKTKLVGFQD-MECVP 137
```

RESULT 2

```
ID   US-08-494-574-7     STANDARD;     PRT;     206 AA.
XX
AC   xxxxxx
XX
DT
DE
XX
```

Sequence 7, Application US/08494574
Patent No. 5783665
GENERAL INFORMATION:

```
CC   APPLICANT: Baum, Peter
CC   APPLICANT: Goodwin, Ray
CC   APPLICANT: Fanslow, William
CC   APPLICANT: Gayle, Richard
CC   TITLE OF INVENTION: No. 5783665el Cytokine Which is a Ligand for
CC   TITLE OF INVENTION: OX40
CC   NUMBER OF SEQUENCES: 13
CC   CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: Immunex Corporation
CC     STREET: 51 University Street
CC     CITY: Seattle
CC     STATE: WA
CC     COUNTRY: USA
CC     ZIP: 98101
CC   COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: IBM PC compatible
CC     OPERATING SYSTEM: PC-DOS/MS-DOS
CC     SOFTWARE: Patentin Release #1.0, Version #1.25
CC   CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: US/08/494,574
CC     FILING DATE: 22-JUN-1995
CC     CLASSIFICATION: 530
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER: US/08/097,827
CC     FILING DATE: 23-JUL-1993
CC   ATTORNEY/AGENT INFORMATION:
CC     NAME: Perkins, Patricia A.
CC     REGISTRATION NUMBER: 34,693
CC     REFERENCE/DOCKET NUMBER: 2806
CC   TELECOMMUNICATION INFORMATION:
CC     TELEPHONE: 206-587-0730
CC   INFORMATION FOR SEQ ID NO: 7:
CC   SEQUENCE CHARACTERISTICS:
CC     LENGTH: 206 amino acids
CC     TYPE: amino acid
CC     TOPOLOGY: linear
CC   MOLECULE TYPE: protein
SQ   SEQUENCE 206 AA; 22939 MW; 232405 CN;

Query Match           5.7%;  Score 178;  DB 1;  Length 206;
Best Local Similarity 27.8%;  Pred. No. 1.81e-05;
Matches   49;  Conservative  36;  Mismatches  77;  Indels  14;  Gaps  11;

Db    9 TALLLIG-LTLGYTARRLNCVKHYPS-GHKC--CRECQPGHGMVNRCD--HTRDTLCHP  62
```

```
Qy   15 TLLVLLGYLSCKVTCESGDCRQEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVA  74
Db   63 CETGFYNEAVNYDfCKQCTQCNHRSGSGSELKONCTPTQDTVC-RCRPGTQPROD-SGYKLG 120
Qy   75 CRLHRFKEDWGFQKCKPCLDCAVVNRFQ-KANCSATSDAICGDCLPGFYRKTKLVGFQ-D 132
Db  121 VDCVPCPPGHFSPGNNQACKPWTNCTLSGKQTRHPASDSLDAV-CED-RSLLATLL 174
Qy  133 MECVPC--GfPPPYEPHCASKVNLVKRIASTASSPRDTAIAAVICSALATVLLALL 186
```

RESULT 3

```
ID   US-08-097-827-7     STANDARD;     PRT;     206 AA.
XX
AC   xxxxxx
XX
DT
DE
XX
```

Sequence 7, Application US/08097827
Patent No. 5457035
GENERAL INFORMATION:

```
CC   APPLICANT: Baum, Peter
CC   APPLICANT: ,Goodwin, Ray
CC   APPLICANT: Fanslow, William
CC   APPLICANT: Gayle, Richard
CC   TITLE OF INVENTION: No. 5457035el Cytokine Which is a Ligand for
CC   TITLE OF INVENTION: OX40
CC   NUMBER OF SEQUENCES: 13
CC   CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: Immunex Corporation
CC     STREET: 51 University Street
CC     CITY: Seattle
CC     STATE: WA
CC     COUNTRY: USA
CC     ZIP: 98101
CC   COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: IBM PC compatible
CC     OPERATING SYSTEM: PC-DOS/MS-DOS
CC     SOFTWARE: Patentin Release #1.0, Version #1.25
CC   CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: US/08/097,827
CC     FILING DATE:
CC     CLASSIFICATION: 536
CC   ATTORNEY/AGENT INFORMATION:
CC     NAME: Perkins, Patricia A.
CC     REGISTRATION NUMBER: 34,693
CC     REFERENCE/DOCKET NUMBER: 2806
CC   TELECOMMUNICATION INFORMATION:
CC     TELEPHONE: 206-587-0730
CC   INFORMATION FOR SEQ ID NO: 7:
CC   SEQUENCE CHARACTERISTICS:
CC     LENGTH: 206 amino acids
CC     TYPE: amino acid
CC     TOPOLOGY: linear
CC   MOLECULE TYPE: protein
SQ   SEQUENCE 206 AA; 22939 MW; 232405 CN;

Query Match           5.7%;  Score 178;  DB 1;  Length 206;
Best Local Similarity 27.8%;  Pred. No. 1.81e-05;
Matches   49;  Conservative  36;  Mismatches  77;  Indels  14;  Gaps  11;

Db    9 TALLLIG-LTLGVTARRLNCVKHYPS-GHKC--CRECQPGHGMVNRCD--HTRDTLCHP  62
Qy   15 TLLVLLGYLSCKVTCESGDCRQOEFRDRSGNCVPCNQCGPGMELSKEGGFGYGEDAQCVA  74
Db   63 CETGFYNEAVNYDfCKQCTQCNHRSGSGSELKQNCTPTQDTVC-RCRPGTQPROD-SGYKLG 120
Qy   75 CRLHRFKEDWGFQKCKPCLDCAVVNRFQ-KANCSATSDAICGDCLPGFYRKTKLVGFQ-D 132
Db  121 VDCVPCPPGHFSPGNNQACKPWTNCTLSGKQTRHPASDSLDAV-CED-RSLLATLL 174
```

Qy 133 MECVPC--GDPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALATVLLALL 186
        ::|||| : :| | :.| |: || | :|| || | | | :|| | || ::| :||

RESULT 4
ID US-08-097-827-11   STANDARD;   PRT;   438 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 11, Application US/0809782T
XX
CC Sequence 11, Application US/0809782T
CC Patent No. 5457035
CC GENERAL INFORMATION:
CC   APPLICANT: Baum, Peter
CC   APPLICANT: Goodwin, Ray
CC   APPLICANT: Fanslow, William
CC   APPLICANT: Gayle, Richard
CC   TITLE OF INVENTION: No. 5457035el Cytokine which is a Ligand for
CC   TITLE OF INVENTION: OX40
CC   NUMBER OF SEQUENCES: 13
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Immunex Corporation
CC   STREET: 51 University Street
CC   CITY: Seattle
CC   STATE: WA
CC   COUNTRY: USA
CC   ZIP: 98101
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC Compatible
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: Patentin Release #1.0, Version #1.25
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/097,827
CC   FILING DATE:
CC   CLASSIFICATION: 536
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Perkins, Patricia A.
CC   REGISTRATION NUMBER: 34,693
CC   REFERENCE/DOCKET NUMBER: 2806
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: 206-587-0730
CC   INFORMATION FOR SEQ ID NO: 11:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 438 amino acids
CC   TYPE: amino acid
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: protein
SQ SEQUENCE   438 AA; 49029 MW; 1097666 CN;

Query Match   5.7%; Score 178; DB 1; Length 438;
Best Local Similarity 27.8%; Pred. No. 1.81e-05;
Matches 49; Conservative 36; Mismatches 77; Indels 14; Gaps 11;

Qy   9 TALLLLG-LTLGVTAARRLNCVKHTYPS-GHKC--CRECQPGHGMVNRCD--HTRDTLCHP 62
         | :|||| |:||| ||| : || | : |: | : ||| || | : |:
Db   9 CETGFYNEAVNYDTCKCQCTQCNHRSGSELKQNCTPTQDTVC-RCRPGTQPROD-SGYKLG 120

Qy   75 CRLHRFKEDWGFQKCKPCLDCAVVNRFQ-KANCSATSDAICGDCLPGFYRKTKLVGFQ-D 132
         ::| ||||| | | : | | |: ||||:|:| | |: ||:| |: || | :| :
Db   63 CETGFYNEAVNYDTCKCQCTQCNHRSGSELKQNCTPTQDTVC-RCRPGTQPROD-SGYKLG 120

Qy 121 VDCVPCPPGHFSPGNNQACKPWTNCTLSGKQTRHPASDSLDAV-CED-RSLLATLL 174
         | :| |||| | | :: | |:| ||:||| |: ::| |: || : :|| |||
Db   75 CRLHRFKEDWGFQKCKPCLDCAVVNRFQ-KANCSATSDAICGDCLPGFYRKTKLVGFQ-D 132

Qy 133 MECVPC--GDPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALATVLLALL 186
Db 121 VDCVPCPPGHFSPGNNQACKPWTNCTLSGKQTRHPASDSLDAV-CED-RSLLATLL 174

RESULT 5
ID US-08-494-574-11   STANDARD;   PRT;   438 AA.
XX

Qy 133 MECVPC--GDPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALATVLLALL 186

--- (second column) ---

AC xxxxxx
XX
DE
DT
XX
Sequence 11, Application US/08494574
CC Sequence 11, Application US/08494574
CC Patent No. 5733665
CC GENERAL INFORMATION:
CC   APPLICANT: Baum, Peter
CC   APPLICANT: Goodwin, Ray
CC   APPLICANT: Fanslow, William
CC   APPLICANT: Gayle, Richard
CC   TITLE OF INVENTION: No. 5783665el Cytokine Which is a Ligand for
CC   TITLE OF INVENTION: OX40
CC   NUMBER OF SEQUENCES: 13
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Immunex Corporation
CC   STREET: 51 University Street
CC   CITY: Seattle
CC   STATE: WA
CC   COUNTRY: USA
CC   ZIP: 98101
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC compatible
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: Patentin Release #1.0, Version #1.25
CC   PRIOR APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/097,827
CC   FILING DATE: 23-JUL-1993
CC   CLASSIFICATION: 530
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/494,574
CC   FILING DATE: 22-JUN-1995
CC   CLASSIFICATION: 530
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Perkins, Patricia A.
CC   REGISTRATION NUMBER: 34,693
CC   REFERENCE/DOCKET NUMBER: 2806
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: 206-587-0730
CC   INFORMATION FOR SEQ ID NO: 11:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 438 amino acids
CC   TYPE: amino acid
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: protein
SQ SEQUENCE   438 AA; 49029 MW; 1097666 CN;

Query Match   5.7%; Score 178; DB 1; Length 438;
Best Local Similarity 27.8%; Pred. No. 1.81e-05;
Matches 49; Conservative 36; Mismatches 77; Indels 14; Gaps 11;

Qy   9 TALLLLG-LTLGVTAARRLNCVKHTYPS-GHKC--CRECQPGHGMVNRCD--HTRDTLCHP 62
         | :|||| |:||| ||| : || | : |: | : ||| || | : |:
Db   9 CETGFYNEAVNYDTCKCQCTQCNHRSGSELKQNCTPTQDTVC-RCRPGTQPROD-SGYKLG 120

Qy   75 CRLHRFKEDWGFQKCKPCLDCAVVNRFQ-KANCSATSDAICGDCLPGFYRKTKLVGFQ-D 132
Db   63 CETGFYNEAVNYDTCKCQCTQCNHRSGSELKQNCTPTQDTVC-RCRPGTQPROD-SGYKLG 120

Qy 121 VDCVPCPPGHFSPGNNQACKPWTNCTLSGKQTRHPASDSLDAV-CED-RSLLATLL 174
Db   75 CRLHRFKEDWGFQKCKPCLDCAVVNRFQ-KANCSATSDAICGDCLPGFYRKTKLVGFQ-D 132

Qy 133 MECVPC--GDPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALATVLLALL 186
Db 121 VDCVPCPPGHFSPGNNQACKPWTNCTLSGKQTRHPASDSLDAV-CED-RSLLATLL 174

RESULT 6
ID US-08-465-982-54   STANDARD;   PRT;   158 AA.
XX
AC xxxxxx
XX
DT
XX

```
DE   Sequence 54, Application US/08465982
XX
CC   Patent No. 5863786
CC
CC   GENERAL INFORMATION:
CC   -APPLICANT: M.Feldmann, P.W. Gray,
CC   APPLICANT: M.J.C. Turner, F.M Brennan
CC   TITLE OF INVENTION: Modified human TNFalpha (Tumor
CC   TITLE OF INVENTION: Necrosis Factor alpha) Receptor
CC   NUMBER OF SEQUENCES: 57
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE:  Reed & Robbins
CC   STREET:  635 Bryant Street
CC   CITY:  Palo Alto
CC   STATE:  California
CC   COUNTRY:  USA
CC   ZIP:  94301
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE:  Floppy disk
CC   COMPUTER:  IBM PC compatible
CC   OPERATING SYSTEM:  PC-DOS/MS-DOS
CC   SOFTWARE:  Patentin Release #1.0, version #1.25
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/050,319
CC   FILING DATE:  10-May-1993
CC   CLASSIFICATION:
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME:  Robbins, Roberta L.
CC   REGISTRATION NUMBER:  33,208
CC   REFERENCE/DOCKET NUMBER:  5150-0030
CC   APPLICATION NUMBER: US/08/465,982
CC   FILING DATE:
CC   PRIOR APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/050,319
CC   FILING DATE:  10-May-1993
CC   CLASSIFICATION:
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE:  (415) 617-8999
CC   TELEFAX:  (415) 327-3231
CC   INFORMATION FOR SEQ ID NO:  54:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH:  158 amino acids
CC   TYPE:  amino acid
CC   TOPOLOGY:  linear
CC   MOLECULE TYPE:  protein
SQ   SEQUENCE 158 AA; 17375 MW; 124033 CN;

Query Match             4.9%;  Score 151;  DB 2; Length 158;
Best Local Similarity   24.6%;  Pred. No. 2.27e-03;
Matches   31;  Conservative  32;  Mismatches 55;  Indels  8;  Gaps  8;

Qy   11 LPL-VLLELLVGIYPSGVIGLLVPHLGDREKRDSV-CPQGKYIHPQNNSICCTKCHKGTYL 68
         ::| ||||  :: :  : :|||  :  : | : :|| :  :|| :  : : :::|:  ::|
Db    1 MALKVLLEQEKTFF-TLLV-LLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMEL 58

Qy   69 YNDCP-GPGQDTDCRECESGSSFTASENHLRHCLSCSKCRKEMGQVEISSCTVDRDTVCT- 126
         : : : : | : | :  : : :: :| | :: :|| : : : :: : :  | : : |:|
Db   59 SKECGFGYGEDAQCVACRLHRFKE-DWGFQKCKPCLDCA-VVNRFQKANCSATSDAICGD 116

Qy  117 CLPGFY 122
         | :||:
Db  127 CHAGFF 132

RESULT 7
ID   US-08-050-319B-54  STANDARD;  PRT;  158 AA.
XX
AC   xxxxxx
XX
DE   Sequence 54, Application US/08050319B
XX
DE   Sequence 54, Application US/08050319B
CC   Patent No. 5633145
CC   GENERAL INFORMATION:
```

```
CC   APPLICANT: M.Feldmann, P.W. Gray,
CC   APPLICANT: M.J.C. Turner, F.M Brennan
CC   TITLE OF INVENTION: Modified human TNFalpha (Tumor
CC   TITLE OF INVENTION: Necrosis Factor alpha) Receptor
CC   NUMBER OF SEQUENCES: 57
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE:  Reed & Robbins
CC   STREET:  635 Bryant Street
CC   CITY:  Palo Alto
CC   STATE:  California
CC   COUNTRY:  USA
CC   ZIP:  94301
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE:  Floppy disk
CC   COMPUTER:  IBM PC compatible
CC   OPERATING SYSTEM:  PC-DOS/MS-DOS
CC   SOFTWARE:  Patentin Release #1.0, version #1.25
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/050,319B
CC   FILING DATE:  10-May-1993
CC   CLASSIFICATION: 435
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME:  Robbins, Roberta L.
CC   REGISTRATION NUMBER:  33,208
CC   REFERENCE/DOCKET NUMBER:  5150-0030
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE:  (415) 617-8999
CC   TELEFAX:  (415) 327-3231
CC   INFORMATION FOR SEQ ID NO:  54:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH:  158 amino acids
CC   TYPE:  amino acid
CC   TOPOLOGY:  linear
CC   MOLECULE TYPE:  protein
SQ   SEQUENCE 158 AA; 17375 MW; 124033 CN;

Query Match             4.9%;  Score 151;  DB 1; Length 158;
Best Local Similarity   24.6%;  Pred. No. 2.27e-03;
Matches   31;  Conservative  32;  Mismatches 55;  Indels  8;  Gaps  8;

Qy   11 LPL-VLLELLVGIYPSGVIGLLVPHLGDREKRDSV-CPQGKYIHPQNNSICCTKCHKGTYL 68
         ::| ||||  :: :  : :|||  :  : | : :|| :  :|| :  : : :::|:  ::|
Db    1 MALKVLLEQEKTFF-TLLV-LLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMEL 58

Qy   69 YNDCP-GPGQDTDCRECESGSSFTASENHLRHCLSCSKCRKEMGQVEISSCTVDRDTVCT- 126
         : : : : | : | :  : : :: :| | :: :|| : : : :: : :  | : : |:|
Db   59 SKECGFGYGEDAQCVACRLHRFKE-DWGFQKCKPCLDCA-VVNRFQKANCSATSDAICGD 116

Qy  117 CLPGFY 122
         | :||:
Db  127 CHAGFF 132

RESULT 8
ID   US-08-219-237B:8  STANDARD;  PRT;  139 AA.
XX
AC   xxxxxx
XX
DT
DE   Sequence 8, Application US/08219237B
XX
CC   Patent No. 5874546
CC   GENERAL INFORMATION:
CC   APPLICANT:  NAGATA, Shigekazu
CC   APPLICANT:  ITOH, Naoto
CC   APPLICANT:  YONEHARA, Shin
CC   TITLE OF INVENTION: DNA Coding for Human Cell Surface Antigen
CC   NUMBER OF SEQUENCES:  11
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE:  James W. Hellwege
CC   STREET:  P.O. Box 2266 Eads Station
```

US-09-490-187-2.rai

Left column

```
CC   CITY: Arlington
CC   STATE: Virginia
CC   COUNTRY: USA
CC   ZIP: 22202
CC   COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: IBM PC compatible
CC     OPERATING SYSTEM: PC-DOS/MS-DOS
CC     SOFTWARE: Patentin Release #1.0, Version #1.25
CC   CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: US 07/872,129
CC     FILING DATE: 28-MAR-1994
CC     CLASSIFICATION: 435
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER: US 07/872,129
CC     FILING DATE: 22-APR-1992
CC     CLASSIFICATION: 435
CC   ATTORNEY/AGENT INFORMATION:
CC     NAME: James W. Hellwege
CC     REGISTRATION NUMBER: 28,808
CC     REFERENCE/DOCKET NUMBER: 516762
CC   INFORMATION FOR SEQ ID NO: 8:
CC     SEQUENCE CHARACTERISTICS:
CC       LENGTH: 139 amino acids
CC       TYPE: amino acid
CC       TOPOLOGY: linear
CC     MOLECULE TYPE: protein
SQ   SEQUENCE 139 AA; 15298 MW; 98625 CN;
```

```
Query Match    4.8%;  Score 148;  DB 2;  Length 139;
Best Local Similarity  29.0%;  Pred. No. 3.84e-03;
Matches  29;  Conservative 20;  Mismatches 45;  Indels  6;  Gaps  4;

Qy  108  ATSDAICGDCLPGTYRKTKLVGFQDMECVPCGDPPPYEP  147

Db   71  PTEDTVC-QCRPGTQPRODSSHKLGVDCVPC--PPGHFSP  107

Db   13  CRECOPGHGMVSRCD--HIRDTVCHPCEPGFYNEAVNYDTCKQCTQCNHRSGSELKQNCT  70

Qy   49  CNOGGPGMELSKEGFGYGEDAQCVACRLHRFKEDWGFQKCKPCLDCAVVNRFQ-KANCS  107
```

```
RESULT  9
ID   US-08-292-549-6     STANDARD;   PRT;   355 AA.
XX
XX   xxxxxx
AC
DE
XX
Sequence 6, Application US/08292549
Patent No. 5464938
GENERAL INFORMATION:
APPLICANT: Smith, Craig A.
APPLICANT: Goodwin, Raymond G.
TITLE OF INVENTION: Isolated Viral Protein TNF Antagonists
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Immunex Corporation
  STREET: 51 University Street
  CITY: Seattle
  STATE: Washington
  COUNTRY: USA
  ZIP: 98101
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/292,549
  FILING DATE:
```

Right column

```
CC     CLASSIFICATION: 530
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER: 07/963,330
CC     FILING DATE: 10/19/92
CC   ATTORNEY/AGENT INFORMATION:
CC     NAME: Perkins, Patricia A.
CC     REGISTRATION NUMBER: 34,693
CC     REFERENCE/DOCKET NUMBER: 2602-A
CC   TELECOMMUNICATION INFORMATION:
CC     TELEPHONE: (206) 587-0430
CC     TELEFAX: (206) 233-0644
CC   INFORMATION FOR SEQ ID NO: 6:
CC     SEQUENCE CHARACTERISTICS:
CC       LENGTH: 355 amino acids
CC       TYPE: amino acid
CC       TOPOLOGY: linear
CC     MOLECULE TYPE: protein
SQ   SEQUENCE 355 AA; 39008 MW; 711992 CN;
```

```
Query Match    4.8%;  Score 148;  DB 1;  Length 355;
Best Local Similarity  29.8%;  Pred. No. 3.84e-03;
Matches  28;  Conservative 19;  Mismatches 40;  Indels  7;  Gaps  6;

Qy   90  KPCLD-CAVVNRFQKANCSATSDAICGDCLPGFY  122

Db   84  LSCNGRCDS-NQVETRSCNTTHNRIC-DCAPGIY  115

Db   27  NGKCKDNEYK-RHHLC--CLSCPPGTYASRLCDSKTNTNIQCTPCASDTFTSRNNHLPAC  83

Qy   31  SGDCRQQEFRDRSGNCGVVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKKE-DWGFQKC  89
```

```
RESULT  10
ID   US-08-232-087A-10     STANDARD;   PRT;   154 AA.
XX
XX   xxxxxx
AC
DE
XX
Sequence 10, Application US/08232087A
Patent No. 5866372
GENERAL INFORMATION:
APPLICANT: Stein, Harald
APPLICANT: Drkop, Horst
APPLICANT: Latza, Ute
TITLE OF INVENTION: Lymphoid CD30-Antigen
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
  STREET: 8110 Gatehouse Road, Suite 500 East
  CITY: Falls Church
  STATE: Virginia
  COUNTRY: U.S.A.
  ZIP: 22042
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/232,087A
  FILING DATE: 08-SEP-1994
  CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
  NAME: Murphy Jr., Gerald M.
  REGISTRATION NUMBER: 28,977
  REFERENCE/DOCKET NUMBER: 756-103P
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (703) 205-8000
  TELEFAX: (703) 205-8050
  TELEX: 248345
```

Left column

```
     INFORMATION FOR SEQ ID NO: 10:
CC     SEQUENCE CHARACTERISTICS:
CC       LENGTH: 154 amino acids
CC       TYPE: amino acid
CC       STRANDEDNESS: not relevant
CC       TOPOLOGY: linear
CC     MOLECULE TYPE: protein
CC     HYPOTHETICAL: NO
CC     FRAGMENT TYPE: Internal
CC     FEATURE:
CC       NAME/KEY: Protein
CC       LOCATION: 1..154
CC       OTHER INFORMATION: /note= "TNFR1, see Fig. 5"
SQ   SEQUENCE 154 AA; 17350 MW; 115236 CN;
```

```
Query Match    4.7%;  Score 147;  DB 2;  Length 154;
Best Local Similarity 26.1%;  Pred. No. 4.58e-03;
Matches  24;  Conservative  20;  Mismatches  44;  Indels  4;  Gaps  4;

Db      2 CPQGKYIHPQNNSICCTKCHKGTYLYNDCP-GPGQDTDCDECESGQSFTASENHLRHCLSC 60
            : ::  : :   :  ::  :: :  ::: :    :::  :::: :    ::|: :::  ||
Qy     34 CRQQEFRDRSGNCVPCNQCGPGMELSKECGPGYGEDAQCVACRLHRFKE-DWGFQKCKPC 92

Db     61 SKCRKEMGQVEISSCTVDRDTVCG-CRKNQYR 91
            :: ::|  :  :: |||:: : ||  :::::
Qy     93 LDCA-VVNRFQKANCSATSDAICGDCLPGFYR 123
```

```
RESULT 11
ID  US-08-050-319B-57    STANDARD;    PRT;    167 AA.
XX
AC  xxxxxx
XX
DE  Sequence 57, Application US/08050319B
```

```
CC   Sequence 57, Application US/08050319B
CC   Patent No. 5633145
CC   GENERAL INFORMATION:
CC     APPLICANT: M.Feldmann, P.W.Gray,
CC     APPLICANT: M.J.C. Turner, F.M Brennan
CC     TITLE OF INVENTION: Modified human TNFalpha (Tumor
CC     TITLE OF INVENTION: Necrosis Factor alpha) Receptor
CC     NUMBER OF SEQUENCES: 57
CC     CORRESPONDENCE ADDRESS:
CC       ADDRESSEE: Reed & Robbins
CC       STREET: 635 Bryant Street
CC       CITY: Palo Alto
CC       STATE: California
CC       COUNTRY: USA
CC       ZIP: 94301
CC     COMPUTER READABLE FORM:
CC       MEDIUM TYPE: Floppy disk
CC       COMPUTER: IBM PC compatible
CC       OPERATING SYSTEM: PC-DOS/MS-DOS
CC       SOFTWARE: Patentin Release #1.0, version #1.25
CC     CURRENT APPLICATION DATA:
CC       APPLICATION NUMBER: US/08/050,319B
CC       FILING DATE: 10-May-1993
CC       CLASSIFICATION: 435
CC     ATTORNEY/AGENT INFORMATION:
CC       NAME: Robbins, Roberta L.
CC       REGISTRATION NUMBER: 33,208
CC       REFERENCE/DOCKET NUMBER: 5150-0030
CC     TELECOMMUNICATION INFORMATION:
CC       TELEPHONE: (415) 617-8999
CC       TELEFAX: (415) 327-3231
CC     INFORMATION FOR SEQ ID NO: 57:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH: 167 amino acids
CC         TYPE: amino acid
CC         TOPOLOGY: linear
CC     MOLECULE TYPE: protein
```

Right column

```
CC   MOLECULE TYPE: protein
SQ   SEQUENCE 167 AA; 18626 MW; 139513 CN;
```

```
Query Match    4.6%;  Score 143;  DB 1;  Length 167;
Best Local Similarity 25.2%;  Pred. No. 9.20e-03;
Matches  32;  Conservative  30;  Mismatches  57;  Indels  8;  Gaps  8;

Qy     11 LPL-VLLELLVGIYPSGVIGLVPHLGDREKRDSV-CPQGKYIHPQNNSICCTKCHKGTYL 68
            ::| |||| |  ::  :  :: :  : :::::    :::::::::::::::::::::::::
Db      1 MALKVLLEQEKTFF-TLLV-LLGILSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMEL 58

Qy     69 YNDCP-GPGQDTDCDECESGQSFTASENHLRHCLSCSKCRKEMGQVEISSCTVDRDTVCG 126
            :::::  |:|:| :   :: ::::: :||:  :|:|  :::  ::  ::||||: ::|:
Db     59 SKECGPGYGEDAQCVACRLHRFKE-DWGFQKCKPCLDCA-VVNRFQKANCSATSDAICGD 116

Qy    127 CRKNQYR 133
            ||  ||
Db    117 CLPGFYR 123
```

```
RESULT 12
ID  US-08-465-982-2    STANDARD;    PRT;    167 AA.
XX
AC  xxxxxx
XX
DT
DE  Sequence 2, Application US/08465982
```

```
CC   Sequence 2, Application US/08465982
CC   Patent No. 5863786
CC   GENERAL INFORMATION:
CC     APPLICANT: M.Feldmann, P.W.Gray,
CC     APPLICANT: M.J.C. Turner, F.M Brennan
CC     TITLE OF INVENTION: Modified human TNFalpha (Tumor
CC     TITLE OF INVENTION: Necrosis Factor alpha) Receptor
CC     NUMBER OF SEQUENCES: 57
CC     CORRESPONDENCE ADDRESS:
CC       ADDRESSEE: Reed & Robbins
CC       STREET: 635 Bryant Street
CC       CITY: Palo Alto
CC       STATE: California
CC       COUNTRY: USA
CC       ZIP: 94301
CC     COMPUTER READABLE FORM:
CC       MEDIUM TYPE: Floppy disk
CC       COMPUTER: IBM PC compatible
CC       OPERATING SYSTEM: PC-DOS/MS-DOS
CC       SOFTWARE: Patentin Release #1.0, version #1.25
CC     PRIOR APPLICATION DATA:
CC       APPLICATION NUMBER: US/08/050,319
CC       FILING DATE: 10-May-1993
CC     CURRENT APPLICATION DATA:
CC       APPLICATION NUMBER: US/08/465,982
CC       FILING DATE:
CC       CLASSIFICATION:
CC     ATTORNEY/AGENT INFORMATION:
CC       NAME: Robbins, Roberta L.
CC       REGISTRATION NUMBER: 33,208
CC       REFERENCE/DOCKET NUMBER: 5150-0030
CC     TELECOMMUNICATION INFORMATION:
CC       TELEPHONE: (415) 617-8999
CC       TELEFAX: (415) 327-3231
CC     INFORMATION FOR SEQ ID NO: 2:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH: 167 amino acids
CC         TYPE: amino acid
CC         TOPOLOGY: linear
CC     MOLECULE TYPE: protein
SQ   SEQUENCE 167 AA; 18626 MW; 139513 CN;
```

```
Query Match    4.6%;  Score 143;  DB 2;  Length 167;
Best Local Similarity 25.2%;  Pred. No. 9.20e-03;
```

Left column

```
Matches   32; Conservative  30; Mismatches  57; Indels   8; Gaps   8;

Qy   11 LPL-VLLELLVGIIPSGVIGLVPHLGDREKRDSV-CPQGKYIHPQNNSICCTKCHKGTYL 68
        :: |||||| ||:::|| ||||| :|:||| |||| |||||| :| ||
Db    1 MALKVLLEQEKTFF-TLLV-LLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMEL 58

Db   69 YNDCP-GPGQDTDCRECESGSSFTASENHLRHCLSCSKCRKEMGQVEISSCTVDRDTVCG- 126
        :: | | |:|| || |:| | |:||| ||| :| | |||
Qy   59 SKECGFGYGEDAQCVACRLHRFKE-DWGFQKCKPCLDCA-VVNRFQKANCSATSDAICGD 116

Db  127 CRKNQYR 133
        || ||
Qy  117 CLPGFYR 123
```

RESULT 13
```
ID   US-08-050-319B-2      STANDARD;      PRT;      167 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 2, Application US/08050319B
XX
CC   Patent No. 5633145
CC   GENERAL INFORMATION:
CC      APPLICANT: M.Feldmann, P.W.Gray,
CC      APPLICANT: M.J.C. Turner, F.M Brennan
CC      TITLE OF INVENTION: Modified human TNFalpha (Tumor
CC      TITLE OF INVENTION: Necrosis Factor alpha) Receptor
CC      NUMBER OF SEQUENCES: 57
CC      CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: Reed & Robbins
CC         STREET: 635 Bryant Street
CC         CITY: Palo Alto
CC         STATE: California
CC         COUNTRY: USA
CC         ZIP: 94301
CC      COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE: Patentin Release #1.0, version #1.25
CC      CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/08/050,319B
CC         FILING DATE: 10-May-1993
CC         CLASSIFICATION: 435
CC      ATTORNEY/AGENT INFORMATION:
CC         NAME: Robbins, Roberta L.
CC         REGISTRATION NUMBER: 33,208
CC         REFERENCE/DOCKET NUMBER: 5150-0030
CC      TELECOMMUNICATION INFORMATION:
CC         TELEPHONE: (415) 617-8999
CC         TELEFAX: (415) 327-3231
CC      INFORMATION FOR SEQ ID NO: 2:
CC         SEQUENCE CHARACTERISTICS:
CC            LENGTH: 167 amino acid
CC            TYPE: amino acids
CC            TOPOLOGY: linear
CC         MOLECULE TYPE: protein
SQ   SEQUENCE 167 AA; 18626 MW; 139513 CN;
```

```
Query Match            4.6%;  Score 143;  DB 1;  Length 167;
Best Local Similarity 25.2%;  Pred. No. 9.20e-03;
Matches   32; Conservative  30; Mismatches  57; Indels   8; Gaps   8;

Db   11 LPL-VLLELLVGIIPSGVIGLVPHLGDREKRDSV-CPQGKYIHPQNNSICCTKCHKGTYL 68
        :: |||||| ||:::|| ||||| :|:||| |||| |||||| :| ||
Qy    1 MALKVLLEQEKTFF-TLLV-LLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMEL 58

Db   69 YNDCP-GPGQDTDCRECESGSSFTASENHLRHCLSCSKCRKEMGQVEISSCTVDRDTVCG- 126
        :: | | |:|| || |:| | |:||| ||| :| | |||
Qy   59 SKECGFGYGEDAQCVACRLHRFKE-DWGFQKCKPCLDCA-VVNRFQKANCSATSDAICGD 116

Db  127 CRKNQYR 133
        || ||
Qy  117 CLPGFYR 123
```

Right column

```
Qy   59 SKECGFGYGEDAQCVACRLHRFKE-DWGFQKCKPCLDCA-VVNRFQKANCSATSDAICGD 116

Db  127 CRKNQYR 133
        || ||
Qy  117 CLPGFYR 123
```

RESULT 14
```
ID   US-08-465-982-57      STANDARD;      PRT;      167 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 57, Application US/08465982
XX
CC   Patent No. 5863786
CC   GENERAL INFORMATION:
CC      APPLICANT: M.Feldmann, P.W.Gray,
CC      APPLICANT: M.J.C. Turner, F.M Brennan
CC      TITLE OF INVENTION: Modified human TNFalpha (Tumor
CC      TITLE OF INVENTION: Necrosis Factor alpha) Receptor
CC      NUMBER OF SEQUENCES: 57
CC      CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: Reed & Robbins
CC         STREET: 635 Bryant Street
CC         CITY: Palo Alto
CC         STATE: California
CC         COUNTRY: USA
CC         ZIP: 94301
CC      COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE: Patentin Release #1.0, version #1.25
CC      CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/08/465,982
CC      PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER: US/08/050,319
CC         FILING DATE: 10-May-1993
CC      ATTORNEY/AGENT INFORMATION:
CC         NAME: Robbins, Roberta L.
CC         REGISTRATION NUMBER: 33,208
CC         REFERENCE/DOCKET NUMBER: 5150-0030
CC      TELECOMMUNICATION INFORMATION:
CC         TELEPHONE: (415) 617-8999
CC         TELEFAX: (415) 327-3231
CC      INFORMATION FOR SEQ ID NO: 57:
CC         SEQUENCE CHARACTERISTICS:
CC            LENGTH: 167 amino acid
CC            TYPE: amino acids
CC            TOPOLOGY: linear
CC         MOLECULE TYPE: protein
SQ   SEQUENCE 167 AA; 18626 MW; 139513 CN;
```

```
Query Match            4.6%;  Score 143;  DB 2;  Length 167;
Best Local Similarity 25.2%;  Pred. No. 9.20e-03;
Matches   32; Conservative  30; Mismatches  57; Indels   8; Gaps   8;

Db   11 LPL-VLLELLVGIIPSGVIGLVPHLGDREKRDSV-CPQGKYIHPQNNSICCTKCHKGTYL 68
        :: |||||| ||:::|| ||||| :|:||| |||| |||||| :| ||
Qy    1 MALKVLLEQEKTFF-TLLV-LLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMEL 58

Db   69 YNDCP-GPGQDTDCRECESGSSFTASENHLRHCLSCSKCRKEMGQVEISSCTVDRDTVCG- 126
        :: | | |:|| || |:| | |:||| ||| :| | |||
Qy   59 SKECGFGYGEDAQCVACRLHRFKE-DWGFQKCKPCLDCA-VVNRFQKANCSATSDAICGD 116

Db  127 CRKNQYR 133
        || ||
Qy  117 CLPGFYR 123
```

```
RESULT    15
ID  US-08-050-319B-25        STANDARD;        PRT;        455 AA.
XX
AC  xxxxxx
XX
DT
XX
DE  Sequence 25, Application US/08050319B
XX
CC  Patent No. 5633145
CC  GENERAL INFORMATION:
CC    APPLICANT: M.Feldmann, P.W. Gray,
CC    APPLICANT: M.J.C. Turner, F.M Brennan
CC    TITLE OF INVENTION: Modified human TNFalpha (Tumor
CC    TITLE OF INVENTION: Necrosis Factor alpha) Receptor
CC    NUMBER OF SEQUENCES: 57
CC    CORRESPONDENCE ADDRESS:
CC    ADDRESSEE:   Reed & Robbins
CC    STREET: 635 Bryant Street
CC    CITY:  Palo Alto
CC    STATE:  California
CC    COUNTRY: USA
CC    ZIP:  94301
CC  COMPUTER READABLE FORM:
CC    MEDIUM TYPE: Floppy disk
CC    COMPUTER: IBM PC compatible
CC    OPERATING SYSTEM: PC-DOS/MS-DOS
CC    SOFTWARE: PatentIn Release #1.0, version #1.25
CC  CURRENT APPLICATION DATA:
CC    APPLICATION NUMBER: US/08/050,319B
CC    FILING DATE: 10-May-1993
CC    CLASSIFICATION: 435
CC  ATTORNEY/AGENT INFORMATION:
CC    NAME: Robbins, Roberta L.
CC    REGISTRATION NUMBER: 33,208
CC    REFERENCE/DOCKET NUMBER: 5150-0030
CC  TELECOMMUNICATION INFORMATION:
CC    TELEPHONE: (415) 617-8999
CC    TELEFAX:  (415) 327-3231
CC  INFORMATION FOR SEQ ID NO: 25:
CC  SEQUENCE CHARACTERISTICS:
CC    LENGTH:  455 amino acids
CC    TYPE:  amino acid
CC    TOPOLOGY:  linear
CC    MOLECULE TYPE:  protein
SQ  SEQUENCE  455 AA;  50579 MW;  1048388 CN;

Query Match            4.6%;             Score 143;  DB 1;  Length 455;
Best Local Similarity  25.2%;            Pred. No. 9.20e-03;
Matches  32;  Conservative  30;  Mismatches  57;  Indels  8;  Gaps  8;

Db   11 LPL-VLLELLVGIYPSGVIGLVPHLGDREKRDSV-CPQGKIIHPQNNSICCTKCHKGTYL 68
         ::| ||||    :||::|:  |: ::|::  :| :|---|-|:::::|  |::||
Qy    1 MALKVLLEQEKTFF-TLJV-LLGYLSCKVTCESGDCRQEFRDRSGNCVPCNQCGPGMEL 58

Db   69 YNDCP-GPGQDTDCRECESGSFTASENHLRHCLSCSKCRKEMGQVEISSCTVDRDTVCG- 126
         :| | |:|||:||| |  : :::||:||: ::|| |:||
Qy   59 SKECGFGYGEDAQCVACRLHRFKE-DWGFQKCKPCLDCA-VVNRFQKANCSATSDAICGD 116

Db  127 CRKNQYR 133
        |   |
Qy  117 CLPGFYR 123
```

Search completed: Tue Apr 18 14:02:57 2000
Job time : 10 secs.

```
M P S R C H                                                         (TM)
```

```
************************************************************************
Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd
************************************************************************
```

Mpsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 18 13:57:47 2000;   MasPar time 22.57 Seconds
                                               443.989 Million cell updates/sec
Tabular output not generated.

Title:
Description: >US-09-490-187-2
                  (1-423) from US09490187.pep
Perfect Score: 3111
Sequence: 1 MALKVLLEQEKTFTLLVLL.........AVIHPATQISLQVRQLGSL 423

Scoring table: PAM 150
                  Gap 11

Searched: 188963 seqs, 23686106 residues
Post-processing: Minimum Match 0%
                  Listing first 45 summaries
Database: a-geneseq36
                  1:geneseqp

Statistics: Mean 34.781; Variance 154.568; scale 0.225

                  Pred. No. is the number of results predicted by chance to have a
                  score greater than or equal to the score of the result being printed,
                  and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description	Pred. No.
1	3111	100.0	423	1	W93581	Human hAPO4-alpha prot	1.36e-283
2	3066	98.6	423	1	W70387	Amino acid sequence of	3.07e-279
3	3040	97.7	417	1	W98146	Human TRAIN-R.	1.01e-276
4	3007	96.7	417	1	W70386	Amino acid sequence of	1.57e-273
5	2177	70.0	416	1	W93579	Mouse mAPO4-alpha (lon	2.24e-193
6	1368	44.0	214	1	W98145	Mouse TRAIN-R (long fo	1.08e-115
7	1368	44.0	214	1	W93580	Mouse mAPO4-alpha (sho	1.08e-115
8	1193	38.3	150	1	W98148	TRAIN-R short, soluble	5.39e-99
9	1008	32.4	150	1	W93583	Mouse mAPO4-gamma prot	2.92e-81
10	1008	32.4	150	1	W98144	Mouse TRAIN-R (short f	2.92e-81
11	328	10.5	109	1	W93582	Rat rAPO4-alpha protei	2.06e-18
12	248	8.0	30	1	W98147	TRAIN-R secreted form	2.94e-11
13	187	6.0	186	1	R62655	Cowpox virus Pst I/Cla	4.09e-06
14	178	5.7	206	1	W48977	Mouse OX40 extracellul	2.23e-05
15	178	5.7	206	1	R81881	Mouse type-II membrane	2.23e-05
16	178	5.7	438	1	R81882	Plasmid pDC406/OX40/Fc	2.23e-05
17	178	5.7	438	1	W48976	OX40/Fc mutein.	2.23e-05
18	172	5.5	95	1	W93584	Mouse mAPO4-beta prote	6.86e-05
19	159	5.1	176	1	W80254	Amino acid sequence of	7.60e-04
20	151	4.9	159	1	R24083	Truncated TNF-alpha 55	3.27e-03
21	150	4.8	181	1	W26708	Human apoptosis protei	3.92e-03
22	148	4.8	355	1	R85073	Cowpox virus T2-equiva	5.63e-03
23	150	4.8	417	1	W95538	Death domain containin	3.92e-03
24	150	4.8	417	1	W26709	Human apoptosis protei	3.92e-03
25	150	4.8	417	1	W31517	Death domain containin	3.92e-03
26	150	4.8	417	1	W57045	Human apoptosis induci	3.92e-03
27	150	4.8	428	1	W55537	Death domain containin	3.92e-03
28	150	4.8	428	1	W31516	Death domain containin	3.92e-03
29	150	4.8	833	1	W64486	Human DR3 protein.	3.92e-03
30	143	4.6	168	1	R24084	Truncated TNF-alpha 55	1.38e-02
31	143	4.6	199	1	R24080	Truncated TNF-alpha 55	1.38e-02
32	143	4.6	211	1	W89225	Tumour necrosis factor	1.38e-02
33	143	4.6	311	1	W89229	Tumour necrosis factor	1.38e-02
34	143	4.6	366	1	W89228	Tumour necrosis factor	1.38e-02
35	143	4.6	371	1	R07449	Tumour Necrosis Factor	1.38e-02
36	143	4.6	397	1	W89227	Tumour necrosis factor	1.38e-02
37	143	4.6	417	1	W89226	Tumour necrosis factor	1.38e-02
38	143	4.6	420	1	W89224	Tumour necrosis factor	1.38e-02
39	143	4.6	443	1	R51033	Tumour necrosis factor	1.38e-02
40	143	4.6	455	1	R11082	Human 55kD TNF-binding	1.38e-02
41	143	4.6	455	1	R42059	Lambda derived TNF-R.	1.38e-02
42	143	4.6	455	1	R31034	Mutant p55 tumour necr	1.38e-02
43	143	4.6	455	1	R10986	30kD TNF inhibitor pre	1.38e-02
44	143	4.6	455	1	R42197	p55 Tumour necrosis fa	1.38e-02
45	143	4.6	455	1	R24000	TNF-alpha 55kD recepto	1.38e-02
```

RESULT
1
W93581 standard; Protein; 423 AA.

ALIGNMENTS

```
ID   W93581 standard; Protein; 423 AA.
AC   W93581;
DT   18-JUN-1999  (first entry)
DE   Human hAPO4-alpha protein.
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW   developmental abnormality; gestational abnormaility; prostate cancer;
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW   apoptosis; human; APO4-alpha.
OS   Homo sapiens.
PN   WO9911791-A2.
PD   11-MAR-1999.
PF   04-SEP-1998; U18393.
PR   05-SEP-1997; US-924634.
PA   (UNIW ) UNIV WASHINGTON.
PI   Chaudhary PM;
PW   WPI; 99-205191/17.
DR   N-PSDB; X23415.
PT   New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT   useful for diagnosis and treatment of prostate cancer and
PT   developmental or gestational abnormalities
PS   Claim 1; Fig 7C; 156pp; English.
CC   This invention describes isolated Tumor Necrosis Factor (TNF) family
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC   their active fragments. APO4 is useful for diagnosing prostate cancer
CC   by determining levels of APO4 in an individual. Prostate cancer can also
CC   be treated using APO4 selective binding agents linked to a therapeutic
CC   moiety. APO4 polypeptides are also useful for identifying selective
CC   binding agents useful in diagnosis/treatment of disease by binding of
CC   agents to the polypeptide/active fragment which is extracelluar, or
CC   expressed on the cell surface. The binding is preferably performed in
CC   vivo. APO4 poly:peptides/ active fragments are also useful for screening
CC   for agonists and antagonists by binding and observing the changer in APO4
CC   activity. Effective pharmacological agents useful in diagnosis or
CC   treatment of d:sease are also identified using APO4 polypeptides/active
CC   fragments and :APO4 signal transducer molecules that specifically interact
CC   with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC   activity. The method is performed in vivo or in vitro. APO polypeptides
CC   are all useful as immunogens for preparing antibodies. APO4 is also
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma cell line
CC   MCF-7, and induced apoptosis.
SQ   Sequence  423 AA;
```

Query Match   100.0%;   Score 3111;   DB 1;   Length 423;

```
RESULT 2

ID   W70387 standard; protein; 423 AA.
AC   W70387;
DT   02-DEC-1998  (first entry)
DE   Amino acid sequence of human beta-OAF065.
OS   Homo sapiens.
KW   cytokine-mediated disease; rheumatism; ulcerative colitis.
KW   Human; beta-OAF065; stroma cell; antibody; Inflammatory;
FH   Key             Location/Qualifiers
FT   Misc_difference 223
FT                   /note= "encoded by AGA"
FT   Misc_difference 224
FT                   /note= "encoded by CCT"
CC   Membrane polypeptide expressed by human stroma cells, and antibodies
CC   recognising it - for treatment of inflammatory and other
CC   cytokine-mediated diseases.
CC   Disclosure: Pages 37-49; 54pp; Japanese.
PS   This is the amino acid sequence of the human beta-OAF065, used in
PT   the method of the invention. The process involves the use of peptides
PT   expressed by stroma cells, and its antibodies are used for in the
PT   prevention and treatment of inflammatory and other cytokine-mediated
PT   diseases such as rheumatism, ulcerative colitis.
SQ   Sequence 423 AA;

PN   WO9838304-A1.
PD   03-SEP-1998.
PD   26-FEB-1998; J00799.
PR   27-FEB-1997; JP-043143.
PA   (ONOY ) ONO PHARM CO LTD.
PI   Fukushima D, Konishi M, Tada H;
DR   WPI; 98-481205/41.
DR   N-PSDB; V33362.
PT   Membrane polypeptide expressed by human stroma cells, and antibodies
```

```
Query Match             98.6%;
Best Local Similarity   98.1%;
Matches   415;  Conservative   5;  Mismatches   3;  Indels   0;  Gaps   0;

Db         1 MALKVLLEQEKTFFLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQGPGMELSK  60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy         1 MALKVLLEQEKTFFLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQGPGMELSK  60
```

```
RESULT 2

ID   W70387 standard; protein; 423 AA.

Best Local Similarity 100.0%;  Pred. No. 1.36e-283;
Matches   423;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Db         1 MALKVLLEQEKTFFLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQGPGMELSK  60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy         1 MALKVLLEQEKTFFLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQGPGMELSK  60

Db        61 ECGFGYGEDAQCVACRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSAISDAICGDCLPG 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        61 ECGFGYGEDAQCVACRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSAISDAICGDCLPG 120

Db       121 FYRKTKLVGFQDMECVPCGDPPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALAT 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       121 FYRKTKLVGFQDMECVPCGDPPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALAT 180

Db       181 VLLALLILCVIYCKRQFMEKKPSWSLRSQDIQYNETELSCFDRPQLHEYAHRACCQCRRD 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       181 VLLALLILCVIYCKRQFMEKKPSWSLRSQDIQYNETELSCFDRPQLHEYAHRACCQCRRD 240

Db       241 SVQTCGPVRLLPSMCCEEACSPNPATLGCGVHSAASLQARNAGPAGEMVPTFFGSLTQSI 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       241 SVQTCGPVRLLPSMCCEEACSPNPATLGCGVHSAASLQARNAGPAGEMVPTFFGSLTQSI 300

Db       301 CGEFSDAWPLMQNPMGGDNISFCDSYPELTGEDIHSLNPELESSTSLDSNSSQDLVGGAV 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       301 CGEFSDAWPLMQNPMGGDNISFCDSYPELTGEDIHSLNPELESSTSLDSNSSQDLVGGAV 360

Db       361 PVQSHSENFTAATDLSRYNNTLVESASTQDALTMRSQLDQESGAVIHPATQTSLQVRQRL 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       361 PVQSHSENFTAATDLSRYNNTLVESASTQDALTMRSQLDQESGAVIHPATQTSLQVRQRL 420

Db       421 GSL 423
             |||
Qy       421 GSL 423
```

```
RESULT 3

ID   W98146 standard; protein; 417 AA.
AC   W98146;
DT   05-JUL-1999  (first entry)
DE   Human TRAIN-R.
KW   TRAIN-R; receptor; human; tumour necrosis factor receptor;
KW   agonist; antagonist; cancer; immunological disease; therapy;
KW   cytostatic.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1..21
FT                   /note= "signal peptide"
FT   Protein         22..417
FT                   /note= "mature protein"
FT   Domain          26..173
FT                   /note= "extracellular domain"
FT   Domain          174..190
FT                   /note= "transmembrane domain"
FT   Domain          191..417
FT                   /note= "cytoplasmic domain"

PN   WO9913078-A1.
PD   18-MAR-1999.
PD   11-SEP-1998; U19030.
PR   06-MAY-1998; US-084422.
PR   12-SEP-1997; US-058631.
PA   (BIOJ ) BIOGEN INC.
PI   Hession C, Tschopp J;
DR   WPI; 99-229238/19.
DR   N-PSDB; X24978.
PT   New cysteine-rich tumor necrosis factor receptor
PS   Claim 2; Page 26; 30pp; English.
CC   The present sequence is a novel human cysteine-rich tumour
CC   necrosis factor receptor family member termed TRAIN-R that is
CC   expressed at low levels in every tissue and cell line tested thus
CC   far, with higher expression detected in heart, prostate, ovary,
CC   testis, peripheral blood lymphocytes, thyroid and adrenal gland.
CC   Cell death can be induced by administering an agent capable of
CC   inhibiting the binding of TRAIN-R to its ligand. A claimed method
CC   of treating, or reducing, the advancement, severity or effects of
CC   an immunological disease in a mammal comprises administering a
CC   pharmaceutical composition which comprises a TRAIN-R blocking agent,
```

```
Db         1 MALKVLLEQEKTFFLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQGPGMELSK  60
```

```
CC     e.g. soluble TRAIN-R. TRAIN-R can be fused to an immunoglobulin to
CC     produce a fusion protein which may be targeted to various sites.
CC     It can be used in binding assays, and to identify antagonists and
CC     agonists. Anti-TRAIN-R antibodies can be used to reduce the
CC     severity of an immune response or to treat cancer. TRAIN-R
CC     blocking agents can also be used to reduce the severity or effects
CC     of an immunological disease (all claimed).
SQ     Sequence 417 AA;

Query Match        97.7%; Score 3040; DB 1; Length 417;
Best Local Similarity  99.3%; Pred. No. 1.0e-276;
Matches  412; Conservative  2; Mismatches  1; Indels  0; Gaps  0;

Qy      1 MALKVLLEQEKTFFILLVLLGYLSCKVTCESGDCRQQEFRDSRSGNCVPCNQCGPGMELSK 60
          ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db      1 MALKVLLEQEKTFFTLLVLLGYLSCKVTCESGDCRQQEFRDSRSGNCVPCNQCGPGMELSK 60

Qy     61 ECGFGYGEDAQCVTCRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSATSDAICGDCLPG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ECGFGYGEDAQCVTCRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSATSDAICGDCLPG 120

Qy    121 FYRKTKLVGFQDMECVPCGDPPPPYYEPHCASKVNLVKIASTASSSPRDTALAAVICSALAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 FYRKTKLVGFQDMECVPCGDPPPPYYEPHCASKVNLVKIASTASSSPRDTALAAVICSALAT 180

Qy    181 VLLALLLLCVIYCKRQFMEKKPSWSLRSQDIQYNETELSCFDRPQLHEYAHRACCQCRRD 240
          ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db    181 VLLALLLLCVIYCKRQFMEKKPSWSLRSQDIQYNGSELSCFDRPQLHEYAHRACCQCRRD 240

Qy    241 SVQTCGPVRLLPSMCCEEACSPNPATLGCGVHSAASLQARNAGPAGEMYPTFFGSLTQSI 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 SVQTCGPVRLLPSMCCEEACSPNPATLGCGVHSAASLQARNAGPAGEMYPTFFGSLTQSI 300

Qy    301 CGEFSDAWPLMQNPMGGDNISFCDSYPELTGEDIHSLNPELESSTSLDSNSSQDLVGGAV 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CGEFSDAWPLMQNPMGGDNISFCDSYPELTGEDIHSLNPELESSTSLDSNSSQDLVGGAV 360

Qy    361 PVQSHSENTAATDLSRYNNTLVESASTQDALTMRSQLDQESGAVIHPATQTSLQ 415
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 PVQSHSENTAATDLSRYNNTLVESASTQDALTMRSQLDQESGAVIHPATQTSLQ 415
```

```
RESULT 4
ID   W70386 standard; Protein; 417 AA.
AC   W70386;
DT   02-DEC-1998 (first entry)
DE   Amino acid sequence of human alpha-OAF065.
OS   Homo sapiens.
KW   cytokine-mediated disease; rheumatism; ulcerative colitis.
KW   Human; alpha-OAF065; stroma cell; antibody; inflammatory;
WPI; 98-481205/41.
DR   N-PSDB; V33361.
DR   WPI; 98-481205/41.
PN   WO983304-A1.
PD   03-SEP-1998;
PF   26-FEB-1998;
PR   27-FEB-1997; JP-043143.
PA   (ONOY ) ONO PHARM CO LTD.
PI   Fukushima D, Konishi M, Tada H;
PT   Membrane polypeptide expressed by human stroma cells, and antibodies
PT   recognising it - for treatment of inflammatory and other
PT   cytokine-mediated diseases.
PS   Claim 1; Pages 28-30; 5pp; Japanese.
FH   Key        Location/Qualifiers
FT   Misc_difference 223
FT               /note= "encoded by AGA""
FT   Misc_difference 224
FT               /note= "encoded by CCT"
CC   This is the amino acid sequence of the human alpha-OAF065, used in
CC   the method of the invention. The process involves the use of peptides
CC   expressed by stroma cells, and its antibodies are used for in the
CC   prevention and treatment of inflammatory and other cytokine-mediated
CC   diseases such as rheumatism, ulcerative colitis.
SQ   Sequence 417 AA;
```

```
Query Match        96.7%; Score 3007; DB 1; Length 417;
Best Local Similarity  98.1%; Pred. No. 1.57e-273;
Matches  407; Conservative  5; Mismatches  3; Indels  0; Gaps  0;

Qy      1 MALKVLLEQEKTFFILLVLLGYLSCKVTCESGDCRQQEFRDSRSGNCVPCNQCGPGMELSK 60
          ||||||||||||||:|||||||||||||:|||||||||||||||||||||||||||||||
Db      1 MALKVLLEQEKTFFLLVLLGYLSCKVTCETGDCRQQEFRDSRSGNCVPCNQCGPGMELSK 60

Qy     61 ECGFGYGEDAQCVTCRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSATSDAICGDCLPG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ECGFGYGEDAQCVACRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSATSDAICGDCLPG 120

Qy    121 FYRKTKLVGFQDMECVPCGDPPPPYYEPHCASKVNLVKIASTASSSPRDTALAAVICSALAT 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 FYRKTKLVGFQDMECVPCGDPPPPYYEPHCASKVNLVKIASTASSSPRDTALAAVICSALAT 180

Qy    181 VLLALLLLCVIYCKRQFMEKKPSWSLRSQDIQYNETELSCFDRPQLHEYAHRACCQCRRD 240
          ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db    181 VLLALLLLCVIYCKRQFMEKKPSWSLRSQDIQYNETELSCFDRPQLHEYAHRACCQCRRD 240

Qy    241 SVQTCGPVRLLPSMCCEEACSPNPATLGCGVHSAASLQARNAGPAGEMYPTFFGSLTQSI 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 SVQTCGPVRLLPSMCCEEACSPNPATLGCGVHSAASLQARNAGPAGEMYPTFFGSLTQSI 300

Qy    301 CGEFSDAWPLMQNPMGGDNISFCDSYPELTGEDIHSLNPELESSTSLDSNSSQDLVGGAV 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db    301 CGEFSDAWPLMQNPMGGDNISFCDSYPELTGEDIHSLNPELESSTSLDSNSSQDLVGGAV 360

Qy    361 PVQSHSENTAATDLSRYNNTLVESASTQDALTMRSQLDQESGAVIHPATQTSLQ 415
          |||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db    361 PVQSHSENFAATDLSRYNNTLVESASTQDALTMRSQLDQESGAVIHPATQTSLQ 415
```

```
RESULT 5
ID   W93579 standard; Protein; 416 AA.
AC   W93579;
DT   18-JUN-1999 (first entry)
DE   Mouse mAPO4-alpha (long) protein.
OS   Mus sp.
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW   developmental abnormality; gestational abnormality; prostate cancer;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW   apoptosis; mouse; mAPO4-alpha.
OS   Mus sp.
PN   WO9911791-A2.
PD   11-MAR-1999.
DT   18-JUN-1999.
PF   04-SEP-1998; U18393.
PR   05-SEP-1997; US-924634.
PA   (UNIW ) UNIV WASHINGTON.
PI   Chaudhary PM;
DR   WPI; 99-205191/17.
DR   N-PSDB; X23413.
PT   New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT   useful for diagnosis and treatment of prostate cancer and
PT   developmental or gestational abnormalities
PS   Claim 1; Fig 7A; 156pp; English.
CC   This invention describes isolated Tumor Necrosis Factor (TNF) family
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC   their active fragments. APO4 is useful for diagnosing prostate cancer
CC   by determining levels of APO4 in an individual. Prostate cancer can also
CC   be treated using APO4 selective binding agents linked to a therapeutic
CC   moiety. APO4 polypeptides are also useful for identifying selective
CC   binding agents useful in diagnosis/treatment of disease by binding of
CC   agents to the polypeptide/active fragment which is extracellular, or
CC   expressed on the cell surface. The binding is preferably performed in
CC   vivo. APO4 polypeptides/active fragments are also useful for screening
CC   for agonists and antagonists by binding and observing the change in
CC   activity. Effective pharmacological agents useful in diagnosis or
CC   treatment of disease are also identified using APO4 polypeptides/active
CC   fragments and APO4 signal transducer molecules that specifically interact
CC   with a cytoplasmic domain of APO4 and detecting a change in level of APO4
```

CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
SQ Sequence 416 AA;

Query Match   70.0%; Score 2177; DB 1; Length 416;
Best Local Similarity 70.5%; Pred. No. 2.24e-193;
Matches 294; Conservative 63; Mismatches 55; Indels 5; Gaps 5;

```
Db      1 MALKVLPLHRTVLFAAILFLLHLACKVSCETTGDCRQQEFKDRSGNCVLCKQCGPGMELSK 60
            ||||||   ||:|: ::||  ||||||||||:||||||||||||||:|||||||||||||||
Qy      1 MALKVLLEQEKTFFILVLLGYLSCKVTCESGDCRQEFRDRSGNCVPCNQCGPGMELSK 60

Db     61 ECGFGYGEDAQCVPCRPHRFKEDWGFQKCKPCADCALVNRFQRANCSHTSDAVCGDCLPG 120
            ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Qy     61 ECGFGYGEDAQCVACRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSAISDAICGDCLPG 120

Db    121 FYRKTKLVGFQDMECVPCGDPPPPYEPHCTSKVNLVKISSTVSSPRDTALAAVICSALAT 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 FYRKTKLVGFQDMECVPCGDPPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALAT 180

Db    181 VLLALLILCVIYCKRQFMEKKPSWSLRSQDIQYNSSELSCFDQDRLRHCAHRACCQYHRD 240
            ||||||||||||||||||||||||||||||||||||||||||||||| |  || : :|
Qy    181 VLLALLILCVIYCKRQFMEKKPS 203
```

CC used to reduce the severity of an immune response or to treat cancer.
CC TRAIN-R blocking agents can be used to reduce the severity or effects
CC of an immunological disease (all claimed).
SQ Sequence 214 AA;

RESULT 6
ID W98145 standard; Protein; 214 AA.
AC W98145;
DT 05-JUL-1999 (first entry)
DE Mouse TRAIN-R (long form).
KW TRAIN-R; receptor; mouse; tumour necrosis factor receptor;
KW agonist; antagonist; cancer; immunological disease; therapy;
OS Mus musculus.
PN WO9913078-A1.
PD 18-MAR-1999.
PF 11-SEP-1998; U19030.
PR 06-MAY-1998; US-084422.
PR 12-SEP-1997; US-058631.
PA (BIOI) BIOGEN INC.
PI Hession C, Tschopp J;
DR WPI; 99-229238/19.
DR N-PSDB; X24977.
PT New cysteine-rich tumor necrosis factor receptor
PS Claim 2; Page 26; 30pp; English.
CC The present sequence is a novel murine cysteine-rich tumour
CC necrosis factor receptor family member termed TRAIN-R (long form).
CC TRAIN-R is expressed at high levels in murine brain and lung, and
CC at lower levels in liver, skeletal muscle and kidney. Cell death
CC can be induced by administering an agent capable of inhibiting the
CC binding of TRAIN-R to its ligand. A claimed method of treating, or
CC reducing, the advancement, severity or effects of an immunological
CC disease in a mammal comprises administering a pharmaceutical
CC composition which comprises a TRAIN-R blocking agent, e.g. soluble
CC TRAIN-R (see W98144). TRAIN-R can be fused to an immunoglobulin
CC molecule to produce a fusion protein which may be targeted to
CC various sites. It can be used in binding assays, and to identify
CC antagonists and agonists. Anti-TRAIN receptor antibodies can be

Query Match   44.0%; Score 1368; DB 1; Length 214;
Best Local Similarity 85.2%; Pred. No. 1.08e-115;
Matches 173; Conservative 17; Mismatches 13; Indels 0; Gaps 0;

```
Db      1 MALKVLPLHRTVLFAAILFLLHLACKVSCETTGDCRQQEFKDRSGNCVLCKQCGPGMELSK 60
            ||||||   ||:|: ::||  ||||||||||:||||||||||||||:|||||||||||||||
Qy      1 MALKVLLEQEKTFFILVLLGYLSCKVTCESGDCRQEFRDRSGNCVPCNQCGPGMELSK 60

Db     61 ECGFGYGEDAQCVPCRPHRFKEDWGFQKCKPCADCALVNRFQRANCSHTSDAVCGDCLPG 120
            |||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Qy     61 ECGFGYGEDAQCVACRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSAISDAICGDCLPG 120

Db    121 FYRKTKLVGFQDMECVPCGDPPPPYEPHCTSKVNLVKISSTVSSPRDTALAAVICSALAT 180
            ||||||||||||||||||||||||||||||| |||||||: |||||||||||||||||||
Qy    121 FYRKTKLVGFQDMECVPCGDPPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALAT 180

Db    181 VLLALLILCVIYCKRQFMEKKPS 203
            |||||||||||||||||||||||
Qy    181 VLLALLILCVIYCKRQFMEKKPS 203
```

RESULT 7
ID W93580 standard; Protein; 214 AA.
AC W93580;
DT 18-JUN-1999 (first entry)
DE Mouse mAPO4-alpha (short) protein.
KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW developmental abnormality; gestational abnormality; prostate cancer;
KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW apoptosis; mouse; mAPO4-alpha.
OS Mus sp.
PN WO9911791-A2.
PD 11-MAR-1999.
PF 04-SEP-1998; U18393.
PF 05-SEP-1997; US-924634.
PA (UNIW) UNIV WASHINGTON.
PR Chaudhary PM;
DR WPI; 99-205191/17.
DR N-PSDB; X23414;
PT New Tumor Necrosis Factor family receptor polypeptides and ligands
PT useful for diagnosis and treatment of prostate cancer and
PS Claim 1; Fig 7B; 156pp; English.
CC This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3)
CC or their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents; useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/ active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the change in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
SQ Sequence 214 AA;

Query Match   44.0%; Score 1368; DB 1; Length 214;

```
Best Local Similarity  85.2%;  Conservative   17;  Pred. No. 1.08e-115;  Mismatches 13;  Indels  0;  Gaps  0;
Matches  173;

Db      1  MALKVLPLHRTVLFAAILFLHLACKVSCETGDCRQQEFKDRSGNCVLCKQCGPGMELSK  60
           ||||||      :::|:::|||::|:|||||||::|||||||||||||||||||||||
Qy      1  MALKVLLEQEKTFFLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMELSK  60

Db     61  ECGFGYGEDAQCVPCRPHRFKEDWGFQKCKPCPCADCALVNRFQRANCSHTSDAVCGDCLPG  120
           ||||||||||:|||||::|| ||||||||:|||||::||||:::||||||||||||||
Qy     61  ECGFGYGEDAQCVACRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSATSDAICGDCLPG  120

Db    121  FYRKTKLVGFQODMECVPCGDPPPPYEPHCTSKVNLVKISSTVSSPRDTALAAVICSALAT  180
           ||||||||||:|||||||||||||||||:||||||||::||||:::||:||||||||
Qy    121  FYRKTKLVGFQODMECVPCGDPPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALAT  180

Db    181  VLLALLICVITCKRQFMEKKPS  203
           ||||||||||||||||||||||
Qy    181  VLLALLICVITCKRQFMEKKPS  203
```

```
RESULT  8
ID     W98148  standard; Peptide; 150 AA.
AC     W98148;
DT     05-JUL-1999  (first entry)
DE     TRAIN-R short, soluble form.
KW     TRAIN-R; receptor; human; tumour necrosis factor receptor;
KW     agonist; antagonist; cancer; immunological disease; therapy;
KW     cytostatic.
OS     Homo sapiens.
PN     WO9913078-A1.
PD     18-MAR-1999.
PF     11-SEP-1998;  U19030.
PR     06-MAY-1998;  US-084422.
PR     12-SEP-1997;  US-058631.
PA     (BIOJ ) BIOGEN INC.
PI     Hession C; Tschopp J;
PI     Browning J L;
DR     WPI; 99-229238/19.
PT     New cysteine-rich tumor necrosis factor receptor
PT     Disclosure: Page 28; 30pp; English.
PS     The present sequence comprises the putative short, secreted soluble
CC     form of a novel human cysteine-rich tumour necrosis factor receptor
CC     family member termed TRAIN-R. The sequence was produced from a
CC     30-amino acid peptide (see W98147) encoded by a cloned exon
CC     sequence (see X24979) and by comparison to the murine TRAIN-R
CC     short form (see W98144). The human soluble TRAIN-R protein is
CC     expected to inhibit signaling by full-length human TRAIN-R (see
CC     W98146). Human TRAIN-R is expressed at low levels in every tissue
CC     and cell line tested thus far, with higher expression detected in
CC     heart, prostate, ovary, testis, peripheral blood lymphocytes,
CC     thyroid and adrenal gland. Cell death can be induced by
CC     administering an agent capable of inhibiting the binding of TRAIN-R
CC     to its ligand. A claimed method of treating, or reducing, the
CC     advancement, severity or effects of an immunological disease in a
CC     mammal comprises administering a pharmaceutical composition which
CC     comprises a TRAIN-R blocking agent, e.g. soluble TRAIN-R. TRAIN-R
CC     can be fused to an immunoglobulin to produce a fusion protein which
CC     may be targeted to various sites. It can be used in binding assays,
CC     and to identify antagonists and agonists. Anti-TRAIN-R antibodies
CC     can be used to reduce the severity of an immune response or to treat
CC     cancer. TRAIN-R blocking agents can also be used to reduce the
CC     severity or effects of an immunological disease (all claimed).
SQ     Sequence  150 AA;
```

```
Query Match             38.3%;  Score 1193;  DB 1;  Length 150;
Best Local Similarity   99.3%;  Pred. No. 5.39e-99;
Matches  148;  Conservative   1;  Mismatches  0;  Indels  0;  Gaps  0;

Db      1  MALKVLLEQEKTFFLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMELSK  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1  MALKVLLEQEKTFFLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMELSK  60

Db     61  ECGFGYGEDAQCVTCRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSATSDAICGDCLPG  120
           ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Qy     61  ECGFGYGEDAQCVACRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSATSDAICGDCLPG  120

Db    121  FYRKTKLVGFQODMECVPCGDPPPPYEPHC  149
           |||||||||||||||||||||||||||||
Qy    121  FYRKTKLVGFQODMECVPCGDPPPPYEPHC  149
```

```
Qy     61  ECGFGYGEDAQCVACRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSATSDAICGDCLPG  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  FYRKTKLVGFQODMECVPCGDPPPPYEPHC  149
           |||||||||||||||||||||||||||||
Qy    121  FYRKTKLVGFQODMECVPCGDPPPPYEPHC  149
```

```
RESULT  9
ID     W93583  standard;  Protein; 150 AA.
AC     W93583;
DT     18-JUN-1999  (first entry)
DE     Mouse mAPO4-gamma protein.
KW     Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW     developmental abnormality; gestational abnormality; prostate cancer;
KW     APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW     cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW     apoptosis; mouse; APO4-gamma.
OS     Mus sp.
PN     WO9911791-A2.
PD     11-MAR-1999.
PF     04-SEP-1998;  U18393.
PR     05-SEP-1997;  US-924634.
PA     (UNIW ) UNIV WASHINGTON.
PI     Chaudhary PM;
DR     WPI; 99-205191/17.
DR     N-PSDB; X23417.
PT     New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT     useful for diagnosis and treatment of prostate cancer and
PT     developmental or gestational abnormalities
PS     Disclosure: Fig 7E; 156pp; English.
CC     This invention describes isolated Tumor Necrosis Factor (TNF) family
CC     receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC     fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC     their active fragments. APO4 is useful for diagnosing prostate cancer
CC     by determining levels of APO4 in an individual. Prostate cancer can also
CC     be treated using APO4 selective binding agents linked to a therapeutic
CC     moiety. APO4 polypeptides are also useful for identifying selective
CC     binding agents. APO4 polypeptides and active fragments are also useful
CC     agents to the polypeptide/active fragment which is extracellular, or
CC     expressed on the cell surface. The binding is preferably performed in
CC     vivo. APO4 polypeptides/ active fragments are also useful for screening
CC     for agonists and antagonists by binding and observing the changer in APO4
CC     activity. Effective pharmacological agents useful in diagnosis or
CC     treatment of disease are also identified using APO4 polypeptides/active
CC     fragments and APO4 signal transducer molecules that specifically interact
CC     with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC     activity. The method is performed in vivo or in vitro. APO polypeptides
CC     are all useful as immunogens for preparing antibodies. APO4 is also
CC     useful for diagnosis/treatment of developmental or gestational
CC     abnormalities. APO8 was transfected to human breast carcinoma cell line
CC     MCF-7, and induced apoptosis.
SQ     Sequence  150 AA;
```

```
Query Match             32.4%;  Score 1008;  DB 1;  Length 150;
Best Local Similarity   81.9%;  Pred. No. 2.06e-81;
Matches  122;  Conservative   15;  Mismatches  12;  Indels  0;  Gaps  0;

Db      1  MALKVLPLHRTVLFAAILFLHLACKVSCETGDCRQQEFKDRSGNCVLCKQCGPGMELSK  60
           ||||||      :|: :::|:||||||||||||||||||||||||| |:|||||||||
Qy      1  MALKVLLEQEKTFFLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMELSK  60

Db     61  ECGFGYGEDAQCVPCRPHRFKEDWGFQKCKPCADCALVNRFQRANCSHTSDAVCGDCLPG  120
           ||||||||||:|||||::|||||||||||||||:||||||||:::|||||||:|||||
Qy     61  ECGFGYGEDAQCVACRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSATSDAICGDCLPG  120

Db    121  FYRKTKLVGFQODMECVPCGDPPPPYEPHC  149
           |||||||||||||||||||||||||||||
Qy    121  FYRKTKLVGFQODMECVPCGDPPPPYEPHC  149
```

```
RESULT  10
ID     W98144  standard;  Protein; 150 AA.
AC     W98144;
```

Left column

DT   05-JUL-1999 (first entry)
DE   Mouse TRAIN-R (short form).
KW   TRAIN-R; receptor; mouse; tumour necrosis factor receptor;
KW   agonist; antagonist; cancer; immunological disease; therapy;
KW   cytostatic.
OS   Mus musculus.
PN   WO9913078-A1
PD   18-MAR-1999.
PF   11-SEP-1998; U19030.
PR   06-MAY-1998; US-084422.
PR   12-SEP-1997; US-058631.
PA   (BIOJ) BIOGEN INC.
PI   Hession C, Tschopp J;
DR   WPI; 99-229238/19.
DR   N-PSDB; X24976.
PT   New cysteine-rich tumor necrosis factor receptor
PS   Claim 2; Page 26; 30pp; English.
CC   The present sequence is a novel murine cysteine-rich tumour
CC   necrosis factor receptor family member termed TRAIN-R.
CC   This putative natural soluble form of murine TRAIN-R (see W98145)
CC   may inhibit signaling by the full-length TRAIN-R. Murine
CC   TRAIN-R is expressed at high levels in brain and lung, and at
CC   lower levels in liver, skeletal muscle and kidney. Cell death
CC   can be induced by administering an agent capable of inhibiting the
CC   binding of TRAIN-R to its ligand. A claimed method of treating, or
CC   reducing, the advancement, severity or effects of an immunological
CC   disease in a mammal comprises administering a pharmaceutical
CC   composition which comprises a TRAIN-R blocking agent, e.g. soluble
CC   TRAIN-R. TRAIN-R can be fused to an immunoglobulin molecule to
CC   produce a fusion protein which may be targeted to various sites.
CC   It can be used in binding assays, and to identify antagonists and
CC   agonists. Anti-TRAIN receptor antibodies can be used to reduce the
CC   severity of an immune response or to treat cancer. TRAIN-R
CC   blocking agents can be used to reduce the severity or effects of an
CC   immunological disease (all claimed).
SQ   Sequence 150 AA;

RESULT 11

ID   W93582 standard; Protein; 109 AA.
AC   W93582;
DT   18-JUN-1999 (first entry)
DE   Rat rAPO4-alpha protein.
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW   developmental abnormality; gestational abnormality; prostate cancer;
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; disease;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW   apoptosis; rat; APO4-alpha.
OS   Rattus sp.
PN   WO9911791-A2.
PD   11-MAR-1999.
PF   04-SEP-1998; U18393.
PR   05-SEP-1997; US-924634.
PA   (UNIW) UNIV WASHINGTON.
PI   Chaudhary PM;
DR   WPI; 99-205191/17.
DR   N-PSDB; X23416.

Query Match   32.48%; Score 1008; DB 1; Length 150;
Best Local Similarity   81.9%;   Pred. No. 2.06e-81;
Matches 122; Conservative 15; Mismatches 12; Indels 0; Gaps 0;

```
Qy    1 MALKVLPLHRTVLFRAILFLHLACKVSCETGDCRQQEERDRSGNCVLKQCGPGMELSK 60
        |||||| |: |:| ::| ||||||| |||||||||||:|:|||||||||:|:||||
Db    1 MALKVLLEQEKTFFTLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMELSK 60

Qy   61 ECGFGYGEDAQCVPCPRHPREKEDWGFQKCKPCADCALVNRFQRANCSHTSDAVCGDCLPG 120
        |||||||||||||||||||||:|||||||||||||||||||||:|||||||||||||||||
Db   61 ECGFGYGEDAQCVACRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSATSDAICGDCLPG 120

Qy  121 FYRKTKLVGFQDMECVPCGDPPPPYEPHC 149
        |||||||||||||||||||||||||||||
Db  121 FYRKTKLVGFQDMECVPCGDPPPPYEPHC 149
```

Right column

PT   New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT   useful for diagnosis and treatment of prostate cancer and
PT   developmental or gestational abnormalities
PS   Claim 1; Fig 7D; 156pp; English.
CC   This invention describes isolated Tumor Necrosis Factor (TNF) family
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC   their active fragments. APO4 is useful for diagnosing prostate cancer
CC   by determining levels of APO4 in an individual. Prostate cancer can also
CC   be treated with APO4 selective binding agents linked to a therapeutic
CC   moiety. APO4 polypeptides are also useful for identifying selective
CC   binding agents; useful in diagnosis/treatment of disease by binding of
CC   agents to the polypeptide/active fragment which is extracellular, or
CC   expressed on the cell surface. The binding is preferably performed in
CC   vivo. APO4 polypeptides/active fragments are also useful for screening
CC   for agonists and antagonists by binding and observing the changer in APO4
CC   activity. Effective pharmacological agents useful in diagnosis or
CC   treatment of disease are also identified using APO4 polypeptides/active
CC   fragments and APO4 signal transducer molecules that specifically interact
CC   with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC   activity. The method is performed in vivo or in vitro. APO polypeptides
CC   are all useful as immunogens for preparing antibodies. APO4 is also
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma cell line
CC   MCF-7, and induced apoptosis.
SQ   Sequence 109 AA;

RESULT 12

ID   W98147 standard; Peptide; 30 AA.
AC   W98147;
DT   05-JUL-1999 (first entry)
DE   TRAIN-R secreted form C-terminal peptide.
KW   TRAIN-R; receptor; human; tumour necrosis factor receptor;
KW   agonist; antagonist; cancer; immunological disease; therapy;
KW   cytostatic.
OS   Homo sapiens.
PN   WO9913078-A1.
PD   18-MAR-1999.
PF   11-SEP-1998; U19030.
PR   06-MAY-1998; US-084422.
PR   12-SEP-1997; US-058631.
PA   (BIOJ) BIOGEN INC.
PI   Hession C, Tschopp J;
DR   WPI; 99-229238/19.
DR   N-PSDB; X24979.
PT   New cysteine-rich tumor necrosis factor receptor
PS   Claim 2; Page 26; 30pp; English.
CC   The present sequence comprises the C-terminal 30 amino acids of a
CC   soluble form of a novel human cysteine-rich tumour necrosis factor
CC   receptor family member termed TRAIN-R. The 30-amino acid peptide
CC   is identical to amino acids 121-149 of the composite human TRAIN-R
CC   protein sequence given in W98146 and to amino acids 121-150 of the
CC   C-terminus of murine TRAIN-R short form (secreted protein, see
CC   W98144). The amino acid sequence of the entire short secreted form
CC   of human TRAIN-R was deduced (see W98148) from the cloned exon
CC   sequence and by comparison to the mouse short form. The human
CC   soluble protein is expected to inhibit signalling by the full-length
CC   TRAIN-R. Human TRAIN-R is expressed at low levels in every tissue
CC   and cell line tested thus far, with higher expression detected in
CC   heart, prostate, ovary, testis, peripheral blood lymphocytes,

Query Match   10.5%; Score 328; DB 1; Length 109;
Best Local Similarity   54.5%;   Pred. No. 2.92e-18;
Matches   60; Conservative 22; Mismatches 23; Indels 5; Gaps 5;

```
Db    1 WPLMQNPLGGDS-SLCDSYPELTGEDINSLNPENESSTSVDSNGGQDLAGAAAP-DS-SG 57
        |||||||:|||: :| ||||||||||||:|||||:|||||:||||| ||:|  :| |
Qy  308 WPLMQNPMGGDNISFCDSYPELTGEDIHSLNPELESSSTSLDSNSSQDLVGGAVPVQSHSE 367

Db   58 DFQENNDSPFYGDADTVWEQTLAQDAQRTPSREGWEASENLNLATSTAFQ 107
        :|:   |||:|||||||:|| ||:|||:|:|||||
Qy  368 NFTAAFDLSRYNNT-LV-ESASTQDALTMRSQLDQESGAVIHPATQTSLQ 415
```

CC thyroid and adrenal gland. Cell death can be induced by
CC administering an agent capable of inhibiting the binding of TRAIN-R
CC to its ligand. A claimed method of treating, or reducing, the
CC advancement, severity or effects of an immunological disease in a
CC mammal comprises administering a pharmaceutical composition which
CC comprises a TRAIN-R blocking agent, e.g. soluble TRAIN-R. TRAIN-R
CC can be fused to an immunoglobulin to produce a fusion protein which
CC may be targeted to various sites. It can be used in binding assays,
CC and to identify antagonists and agonists. Anti-TRAIN-R antibodies
CC can be used to reduce the severity of an immune response or to treat
CC cancer. TRAIN-R blocking agents can also be used to reduce the
CC severity or effects of an immunological disease (all claimed).
SQ Sequence 30 AA;

Query Match   8.0%;
Best Local Similarity 100.0%;   Score 248; DB 1; Length 30;
Matches   29; Conservative   0; Mismatches   0; Indels   0; Gaps   0;   Pred. No. 2.94e-11;

Qy   1 FYRKTKLVGFQDMECVPCGDPPPPYEPHC 29
          |||||||||||||||||||||||||||||
Db   121 FYRKTKLVGFQDMECVPCGDPPPPYEPHC 149

RESULT 13
ID   R62655 standard; Protein; 186 AA.
AC   R62655;
DT   30-JUN-1995 (first entry)
DE   Cowpox virus fragment Pst I/Cla I fragment viral protein.
KW   Cowpox virus fragment Pst I/Cla I; cytokine antagonist;
KW   tumour necrosis factor; immune response regulation.
OS   Cowpox virus.
PN   US535039-A.
PD   25-OCT-1994.
PR   09-JUL-1993; 089458.
PR   09-JUL-1993; US-089458.
PA   (IMMV) IMMUNEX CORP.
PI   Goodwin RG, Smith CA;
DR   WPI; 94-341063/42.
DR   N-PSDB; Q72995.
CC   Isolated viral proteins capable of binding TNF - therefore
CC   functioning as TNF and cytokine antagonists
PT   Claim 1; Columns 21-22; 13pp; English.
PS   Q72995 encodes R62655 a viral protein from the Pst I/Cla I
CC   fragment of the cowpox virus, which binds cytokines and tumour
CC   necrosis factor (TNF). The viral protein can be used to regulate
CC   immune response as part of a therapeutic composition, it can also
CC   be used as an antagonist of TNF.
SQ   Sequence 186 AA;

Query Match   6.0%;
Best Local Similarity 30.0%;   Score 187; DB 1; Length 186;
Matches   33; Conservative 28; Mismatches 40; Indels   9; Gaps   9;   Pred. No. 4.09e-06;

Db   34 NGSCDDGEYLDKTHN-QCCNRCPPG-EFAKIRCS-G-SDNTKCERCPPHYTTVPNYSNG 89
Qy   31 SGDCRQEFRDSRGNCVPCNOCGPGMELSKE-CGFGYGEDAQCVACRLHRFEKEDWGFQK- 88

Db   90 CHQCRKCPT-GSFDKVKCTGTQNSKCS-CLPGWFCATDSSKTEDCRDCIP 137
Qy   89 CKPCLDCAVVNRFQKANCSATSDAICGDCLPGFYRKTKLVGFQD-MECVP 137

RESULT 14
ID   W48977 standard; Protein; 206 AA.
AC   W48977;
DT   25-SEP-1998 (first entry)
DE   Mouse OX40 extracellular domain.
KW   OX40; cytokine; T cell antigen; TH-2 immune response; OX40-L;
KW   Mus sp.
PN   US5783665-A.
PD   21-JUL-1998.
PF   22-JUN-1995; 494574.
PR   23-JUL-1993; US-097827.
PR   22-JUN-1995; US-494574.
PA   (IMMV) IMMUNEX CORP.
PI   Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
DR   WPI; 98-427099/36.
DR   N-PSDB; V32640.
PT   Purified polypeptide OX-40 ligands -- for co-stimulation of T-cell
PT   production and binding assays for OX-40 and homologues
PS   Example 1; Col 33-34; 26pp; English.
CC   The present sequence represents the mouse OX40 extracellular domain
CC   The extracellular domain of OX40 is its ligand binding domain. The
CC   cDNA (V32640) encoding OX40 extracellular domain was used in the
CC   construction of the chimeric OX40/FC cDNA (V32640). The invention
CC   claims for a murine OX40-L cytokine (W48975) that binds to the OX40
CC   murine T cell antigen. The OX40-L protein is claimed to be useful
CC   for co-stimulation of T-cell production and in binding assays for
CC   detecting OX40 or its homologues. The OX40-L protein is also claimed
CC   to generate a TH-2 immune response.
SQ   Sequence 206 AA;

Query Match   5.7%;
Best Local Similarity 27.8%;   Score 178; DB 1; Length 206;
Matches   49; Conservative 36; Mismatches 77; Indels 14; Gaps 11;   Pred. No. 2.23e-05;

Db   9 TALLLLG-LTLGVTARRLNCVKHTYPS-GHKC--CRECQPGHGMVNRCD--HTRDTLCHP 62
Qy   15 TLLVLLGYLSCKVTCESGDCRQEFRDSRGNCVPCNOCGPGMELSKECEGFGYGEDAQCVA 74

Db   63 CETGFYNEAVNDTCKQCTQCNHRSGSSELKQNCTPTQDTVC-RCRPGTQPROD-SGYKIG 120
Qy   75 CRLHRFEKEDWGFQKCKPCLDCAVVNRFQ-KANCSATSDAICGDCLPGFYRKTKLVGFQ-D 132

Db   121 VDCVPCPPGHFSPGNNQACKPWTNCTLSGKQTRHPASDSLDAV-CED-RSILATLL 174
Qy   133 MECVPC--GDPPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALATVLIALL 186

RESULT 15
ID   R81881 standard; Protein; 206 AA.
AC   R81881;
DT   08-JUL-1996 (first entry)
DE   Mouse type-II membrane polypeptide OX40 extracellular domain.
KW   OX40; OX40-L; cytokine; cell surface molecule;
KW   membrane glycoprotein.
OS   Mus musculus.
PN   US5457035-A.
PD   10-OCT-1995.
PR   23-JUL-1993; 097827.
PR   23-JUL-1993; US-097827.
PA   (IMMV) IMMUNEX CORP.
PI   Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
DR   WPI; 95-357992/46.
DR   N-PSDB; T00826.
PT   New isolated DNA encoding the OX40 ligand polypeptide - also vectors
PT   and host cells used to produce recombinant ligand used in e.g.
PT   prim. T cell culture, to modulate immune response etc.
PS   Example 1; Column 33-34; 26pp; English.
CC   This sequence encodes the extracellular domain of OX40, a membrane
CC   glycoprotein present on the CD4 positive subset of activated T
CC   cells.
SQ   Sequence 206 AA;

Query Match   5.7%;
Best Local Similarity 27.8%;   Score 178; DB 1; Length 206;
Matches   49; Conservative 36; Mismatches 77; Indels 14; Gaps 11;   Pred. No. 2.23e-05;

Db   9 TALLLLG-LTLGVTARRLNCVKHTYPS-GHKC--CRECQPGHGMVNRCD--HTRDTLCHP 62
Qy   15 TLLVLLGYLSCKVTCESGDCRQEFRDSRGNCVPCNOCGPGMELSKECEGFGYGEDAQCVA 74

Db   63 CETGFYNEAVNDTCKQCTQCNHRSGGSELKQNCTPTQDTVC-RCRPGTQPROD-SGYKIG 120
Qy   75 CRLHRFEKEDWGFQKCKPCLDCAVVNRFQ-KANCSATSDAICGDCLPGFYRKTKLVGFQ-D 132

```
Db    121 VDCVPCPPGHFSPGNNQACKPWTNCTLSGKQTRHPASDSLDAV-CED-RSLLATLL 174
          ::|||| | :| : | :| | | :|:|| :|| :||
Qy    133 MECVPC--GDPPPYEPHCASKVNLVKIASTASSPRDTALAAVICSALATVLLALL 186
```

Search completed: Tue Apr 18 13:58:13 2000
Job time : 26 secs.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2000, 01:27:45 ; Search time 463.49 Seconds
(without alignments)
12129.633 Million cell updates/sec

Title: US-09-490-187-1
Perfect score: 1489
Sequence: 1 ggacctgcagcctccaggt.........gtattctttaaaaaacttt 1489

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4538634 seqs, 1887831982 residues

Total number of hits satisfying chosen parameters: 9077268

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:*

```
 1: em_est1:*       45: gb_est26:*
 2: em_est2:*       46: gb_est27:*
 3: em_est3:*       47: gb_est28:*
 4: em_est4:*       48: gb_est29:*
 5: em_est5:*       49: gb_est30:*
 6: em_est6:*       50: gb_est31:*
 7: em_est7:*       51: gb_est32:*
 8: em_est8:*       52: em_est20:*
 9: em_est9:*       53: em_est21:*
10: em_est10:*      54: em_est22:*
11: em_est11:*      55: em_est23:*
12: em_est12:*      56: em_est24:*
13: em_est13:*      57: em_est25:*
14: em_est14:*      58: em_est26:*
15: em_est15:*      59: gb_est33:*
16: em_est16:*      60: gb_est34:*
17: em_est17:*      61: gb_est35:*
18: em_est18:*      62: gb_est36:*
19: em_est19:*      63: gb_est37:*
20: gb_est1:*       64: gb_est38:*
21: gb_est2:*       65: em_est27:*
22: gb_est3:*       66: em_est28:*
23: gb_est4:*       67: em_est29:*
24: gb_est5:*       68: em_est30:*
25: gb_est6:*       69: gb_est39:*
26: gb_est7:*       70: gb_est40:*
27: gb_est8:*       71: gb_est41:*
28: gb_est9:*       72: gb_est42:*
29: gb_est10:*      73: gb_est43:*
30: gb_est11:*      74: gb_est44:*
31: gb_est12:*      75: em_est31:*
32: gb_est13:*      76: em_est32:*
33: gb_est14:*      77: em_est33:*
34: gb_est15:*      78: em_est34:*
35: gb_est16:*      79: gb_gss1:*
36: gb_est17:*      80: gb_gss2:*
37: gb_est18:*      81: gb_gss3:*
38: gb_est19:*      82: gb_gss4:*
39: gb_est20:*      83: em_gss1:*
40: gb_est21:*      84: em_gss2:*
41: gb_est22:*      85: em_gss3:*
42: gb_est23:*      86: gb_gss5:*
43: gb_est24:*      87: gb_gss6:*
44: gb_est25:*      88: gb_gss7:*
                    89: gb_gss8:*
                    90: gb_gss9:*
                    91: em_gss4:*
                    92: em_gss5:*
                    93: em_gss6:*
                    94: em_gss7:*
                    95: em_gss8:*
                    96: em_gss9:*
                    97: em_gss10:*
                    98: em_gss11:*
                    99: gb_gss10:*
                   100: gb_gss11:*
                   101: em_gss12:*
                   102: gb_gss12:*
                   103: gb_gss13:*
                   104: gb_gss14:*
                   105: gb_gss15:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match %	Length	DB	ID	Description
1	454	30.5	454	26	W56629	W56629 zd16e11.r1
2	272.2	18.3	282	61	AL036000	AL036000 DKFZp564K

```
RESULT 1

3    214     14.4    401    27   AA003356    AA003356 mg49g01.r
4    185     12.4    358    27   AA036247    AA036247 mi74a03.r
5    153.4   10.3    643   104   AQ563354    AQ563354 HS_5335_B
6    124.4    8.4    275    59   AV111112    AV111112 AV111112
7    122.6    8.2    646   104   AQ544065    AQ544065 RPCI-11-3
8    116      7.8    404    34   AA495217    AA495217 fa04d10.r
9    112.2    7.5    450   104   AA514075    AA514075 HS_5184_A
10   78.6     5.3    381    48   AI551729    AI551729 vf11e02.y
11   61.8     4.2    485    64   AL120773    AL120773 DKF2p762C
12   59       4.0    405    99   AQ173919    AQ173919 HS_3204_A
13   43.4     2.9   1101    79   CNS00LPM    CNS00LPM Drosophil
14   42.6     2.8    443    38   AA759377    AA759377 ah54a10.s
15   42.4     2.8    806    79   CNS009YZ    CNS009YZ Drosophil
16   41.4     2.8    551    38   AA814296    AA814296 nz07d08.s
17   41.2     2.8   1101    79   CNS00K81    CNS00K81 Drosophil
18   40.6     2.7    313    79   AI668861    AI668861 wb82b04.x
19   40.6     2.7   1101    79   CNS00BEU    CNS00BEU Drosophil
20   40.4     2.7   1201    80   CNS01G2K    CNS01G2K Drosophil
21   40.2     2.7    697    80   CNS018M5    CNS018M5 Drosophil
22   40.2     2.7    778    38   AA748343    AA748343 oa55b03.s
23   40       2.7    209    74   AW196030    AW196030 x186f10.x
24   40       2.7    948    80   CNS0159T    CNS0159T Drosophil
25   39.8     2.7   1101    79   CNS00LJT    CNS00LJT Drosophil
26   39.8     2.7    242    73   AW168427    AW168427 xi83c09.x
27   39.6     2.7    494    74   AW194174    AW194174 xm07c04.x
28   39.6     2.7   1009    79   CNS00587    CNS00587 Drosophil
29   39.6     2.7   1047    80   CNS0164P    CNS0164P Drosophil
30   39.6     2.7    397    39   AA860208    AA860208 ak48c10.s
31   39.4     2.6    465    26   W47341      W47341 zc34h02.s1
32   39.4     2.6    962    79   CNS007P9    CNS007P9 Drosophil
33   39.4     2.6   1101    79   CNS00M7T    CNS00M7T Drosophil
34   39.2     2.6    418    63   AI963611    AI963611 wr65e05.x
35   39.2     2.6    540    48   AI610254    AI610254 tp14a09.x
36   39       2.6    207    44   AU037991    AU037991 AU037991
37   39       2.6    407    26   W32428      W32428 zc05c12.s1
38   39       2.6    448    81   B51424      B51424 CIT978SK-18
39   39       2.6    559    29   AA161105    AA161105 zo58c05.s
40   39       2.6    964    80   CNS017VX    CNS017VX Drosophil
41   39       2.6   1101    79   CNS0022U    CNS0022U Drosophil
42   38.8     2.6    925    79   CNS0091P    CNS0091P Drosophil
43   38.8     2.6   1101    79   CNS00AY1    CNS00AY1 Drosophil
44   38.6     2.6    539    44   AI258473    AI258473 LP01680.5
45   38.6     2.6    594    43   AU034113    AU034113 AU034113
```

```
RESULT 1

LOCUS       W56629        454 bp    mRNA                       EST          15-OCT-1996
DEFINITION  zd16e11.r1 Soares_fetal_heart_NbHH19w Homo sapiens cDNA clone
            IMAGE:340844 5', mRNA sequence.
ACCESSION   W56629
VERSION     W56629.1  GI:1358487
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 454)
AUTHORS     Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
TITLE       The WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT     On Apr 14, 1993 this sequence version replaced gi:785895.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: mob.REGA+ET.
            Insert Length: 1373   Std Error: 0.00
FEATURES             Location/Qualifiers
     source          1..454
                     /organism="Homo sapiens"
                     /db_xref="GDB:1266219"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:340844"
                     /clone_lib="Soares_fetal_heart_NbHH19w"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGGCGCCGCCATCGTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library was constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus) as the fetal lung library, Soares fetal lung
                     NbHH19W."
BASE COUNT      132 a     96 c     88 g    138 t
ORIGIN
```

```
ALIGNMENTS

Query Match           454; Best Local Similarity   100.0%;
Best Local Similarity   30.5%;   Score 454;   DB 26;   Length 454;
Matches    454;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

                        Score 454;   Pred. No. 2.8e-109;

Qy   121   TCATTCTGAAAACTTTACAGCAGCTACTTGATTTATCTAGATATAACAACACACTGGTAGA 180
           ||||||||||||||| |||||||||||||||||| |||||||||||||||||||| |||||
Db  1156   tcatcctgaaaacttacagcagctacttgattttatctagatataacaacacacggtaga 1215

Qy   181   ATCAGCATCAACTCAGGATCGCACTCAACTATGAGAAGAAGAGCCACCTAGATCAGAGAGACTGGCGC 240
           |||||||||||||||||||||||||||| |||||||||||||||||||||| ||||||||||||||||
Db  1216   atcagcatcaactcaggatcgcactcaactcaactatgagaagaagagccagctagatcagagagactggcgc 1275

Qy   241   TGTCAATCCACCAGCCACTCAGAGCGTCCCTCCAGGTAAGCAGCAGCCACTGGGTTCCCTGTG 300
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1276   tgtcatccaccagccactcagacgacgtccctccaggtaagcagcagccactggggttccctgtg 1335

Qy   301   AACACAGCACTGACTTACAGTAGCAGTAGATCAAGAACTCTGTTCCCAGCATAAGATTTGGGGGGAAC 360
           ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  1336   aacacagcacactgacttacagtagcagtagatcaagaactctgttcccagcataagatttgggggaac 1395

Qy   361   CTGGATGGAGTTTTTTTTTTTTGCATGTTTATTAATTTCTTATTATGTTGTAGAGTATGTTTT 420
           ||||||||||||||| ||||||| ||||||||||||||||||||||||||| |||||||||
Db  1396   ctggatgagttttttttttttgcatcttaataattctttataatgttgtagagtatgttttt 1455

Qy   421   AAAATAAATTCAAGTATTTTTTTTTAAAAACTTT 454
           ||||||||||||||||||||||||||||||||||
Db  1456   aaaataaattcaagtatttttttttaaaaaactt 1489
```

```
RESULT 2

LOCUS       AL036000       282 bp    mRNA                       EST          27-SEP-1999
DEFINITION  DKF2p564K1022_r1 564 (synonym: hfbr2) Homo sapiens cDNA clone
```

```
ACCESSION   DKFZp564K1022 5', mRNA sequence.
VERSION     AL036000
KEYWORDS     AL036000.1    GI:5405629
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 282)
AUTHORS     Wambutt,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE       EST (Wambutt, et al.)
JOURNAL     Unpublished (1999)
COMMENT     On Jun 22, 1998 this sequence version replaced gi:3246658.
            Contact: Wambutt R
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by AGOWA within the cDNA sequencing consortium of the
            German Genome Project.
            si sequence also available.
            This clone is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1..282
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp564K1022"
                     /clone_lib="564 (synonym: hfbr2)"
                     /tissue_type="brain"
                     /dev_stage="fetal"
                     /lab_host="X1-2blue"
                     /note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"
BASE COUNT       57 a       85 c       75 g       63 t       2 others
ORIGIN

Query Match         18.3%;    Score 272.2; DB 61;  Length 282;
Best Local Similarity     98.2%;     Pred. No. 2.4e-61;
Matches  274;  Conservative    0;  Mismatches   5;  Indels    0;  Gaps    0;

Qy     429 ataggaagacgaaacttgtcggcttcaagacactaggagtgtgtccttgtggagaccctc 488
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       3 ATAGGAAGACGAAACTTGTCGGCTTTCAAGACACAGGAGTGTGTGCCTTGTGGAGACCCTC 62

Qy     489 ctcctcttacgaaccgcaactgtgccagcaagtcaacctctgaagatcgcgtccacgg 548
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 CTCCTCCTTACGAACCGCACTGTGCCAGCAAGGTCAACCTCGTGAAGATCGCGTCCACGG 122

Qy     549 cccccagccacgggacacggcgcgctggcctgccgttatctgcaggcgctctggccacgtcc 608
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 CCTCCAGCCACCACGGGACACGGCGCGCTGGCGTCCCGTTATCTGCAGCGCTCGGCCACCGTCC 182

Qy     609 tgctggcgcctgctcatcctctgtgtcattcattgtaaggacagttatggaaaagaaac 668
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 TGCTGGCGCCCTGCCTCATCCCTCTGTGTCATCTATTTGTAAGGACAGTTTATGGAGAAGAAAC 242

Qy     669 ccagctgtctctgcgtcacaggacaattcagtacaacg 707
           ||||||||||||||||||||||||||||||||||||||
Db     243 CCACCTGGTCTCTCCGGTGCCANGACAATTCAGTACTACG 281
```

```
RESULT   3
LOCUS       AA003356
ACCESSION   AA003356
DEFINITION  mg49g01.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:427152 5', mRNA sequence.
ACCESSION   AA003356
VERSION     AA003356.1    GI:1446796
KEYWORDS    EST.
SOURCE      house mouse.
LOCUS       AA003356        401 bp    mRNA        EST       19-JUL-1996
DEFINITION  mg49g01.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:427152 5', mRNA sequence.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 401)
AUTHORS     Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,N., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Thelsing,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE       The WashU-HHMI Mouse EST Project
JOURNAL     Unpublished (1996)
COMMENT     On May 8, 1995 this sequence version replaced gi:799469.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314-286 1800
            Fax: 314 286 1810
            Email: mousestt@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:261/04
            Seq primer: ETprimer
            High quality sequence stop: 345.
FEATURES             Location/Qualifiers
     source          1..401
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:427152"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGACGGGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
                     State Univ., from 2 j; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo."
BASE COUNT       92 a       96 c      128 g       85 t
ORIGIN

Query Match         14.4%;    Score 214; DB 27;  Length 401;
Best Local Similarity     77.5%;     Pred. No. 5.8e-46;
Matches  259;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy      38 caacaataacatacaatttgataagaaagaatggcttaaaagtgctactagaacaagagaa 97
           ||   ||||||||| | |||||||||||||||||| |||||||||||||||||
Db      68 CAGGAATAAACACGTTTGGTGAGAACCCATGGCACTCAAGGTCCTACCTCTACACAGGACG 127

Qy      98 acgttttctcctctttttagttattactaggctattgtcatgtaaagtgactgtgaatca 157
           || ||||||||||||| |||||||||||||||||||  ||||||| || ||||||||
Db     128 GTGCCTGTTCCGTGCCATTCTCTTCCTTACTCCACCTGGCAGTGAAAGTGAGGTTGCGAAAACC 187

Qy     158 gggagactgtagacaacaagaaatttcagggatcggtctggaaactgttgttccctgcaacag 217
           || |||| ||  ||||||||||||||  ||||||||||||| |||||||  ||||| |||
Db     128 GTGCCTGTTCCGTGCCATTCTCTTCCTTACTCCACCTGGCAGTGAAAGTGAGGTTGCGAAAACC 187

Qy     218 tgtgggccagtcatggagttgtctcaagggaatgtgcttccgcctatgggggaggatcgcaag 277
           ||   ||||||||| | ||| |||||||||||||| || |||||| ||||||||||||||
Db     248 TGCGGACCTGCCATGGAGTTGTCTCCAAGGAATGTGCTTCCGCCTATGGGGAGGATCCACAG 307

Qy     278 tgtgtggcgcgcggctgcacaggttcaaggggactgggggcttccagaaatgcaagccc 337
           |||| || ||| | ||| ||| | | |||||| |||||| || || | | |
```

```
RESULT    4

LOCUS       AA036247                358 bp    mRNA                    EST       26-AUG-1996
DEFINITION  ml74a03.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
            IMAGE:472300 5', mRNA sequence.
ACCESSION   AA036247
VERSION     AA036247.1  GI:1509376
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 358)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On Apr 14, 1993 this sequence version replaced gi:716824.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:283044
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 347.
FEATURES             Location/Qualifiers
     source          1..358
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:472300"
                     /clone_lib="Soares mouse p3NMF19.5"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGAGGCCGGAGCCGCCATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA was kindly provided by
                     Dr. Minoru Ko (Wayne State University)."
BASE COUNT       81 a    98 c    110 g    69 t
ORIGIN

Query Match        12.4%;  Score 185;  DB 27;  Length 358;
Best Local Similarity   77.5%;
Matches  224;     Conservative   0;     Mismatches  65;     Indels   0;     Gaps   0;

Qy          42 aataaatatacatttgataaagaagatggctttaaaagtgctactagaacaagagaaacgt 101
               ||||||||||||||||| ||||||| |||||||| |||| |||| || ||||
Db          69 AATAAAACACGTTTGGTAGAGAACCCATGGCACTCAAGGTCCTTACCCTTACACAGGAGTC 128
Qy         102 ttttcacttctttagtattactagagctatttgtcatgtaaagtgacttgtgaatacagga 161
               |||||| || || | || | || ||| | |||| ||
```

```
RESULT    5

LOCUS       AQ563354                643 bp    DNA                     GSS       29-MAY-1999
DEFINITION  HS_5335_B2_B03_T7A RPCI-11 Human Male BAC library Homo sapiens
            genomic clone Plate=911 Col=6 Row=D, genomic survey sequence.
ACCESSION   AQ563354
VERSION     AQ563354.1  GI:4922825
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 643)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased
            or from Research Genetics (info@resgen.com). BAC end Web Server:
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            http://www.htsc.washington.edu
FEATURES             Location/Qualifiers
     source          1..643
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="plate=911 Col=6 Row-D"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Genomic sequence of BAC ends"
            Plate: 911 row: D column: 6
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 643.
BASE COUNT      151 a   154 c   158 g   167 t   13 others
ORIGIN

Query Match        10.3%;  Score 153.4;  DB 104;  Length 643;
Best Local Similarity   85.4%;
Matches  169;     Conservative   0;     Mismatches  29;     Indels   0;     Gaps   0;

Qy         509 tgtgccagcaagttcaacctcgtgaagatcgcgtccacacgcgcctccagccacgggdacacg 568
               |||||||||||||| || || ||||| ||| | ||| || ||| || |||||| |||||
Db         315 TGTGCCAGCAAGGTCAACCTCGTGAAGATGCGCTNGCACGGCCTCCAGCCCACGGGACACG 374
Qy         569 gcgctggctgccgttatctcggccaggccctggccacgtcctgctggccgctcatcctc 628
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         308 TGTGTGCCCTGCAGGCGCGCCGCCACCGGTTCAAGGAAGACTGGGGGTTTCAAGAAGTGTAAGCCA 367
               |||||| |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Qy         338 tgtctggaactgcgcgcaatgggtgaaccgcttcaga 371
               ||||||| || ||||| | | ||||||||| |||||
Db         368 TGTGCGGACTGCGCGCTGTGAACCGCTGTCAGA 401
```

```
Db         129 TCTTCGTTGCCTTGCCATTTCTTCCTTACTCCACCTGGCATTGTAAAGTGAGTTGCGAAACCGGAG 188
               ||| ||||||| ||| ||| |||| |||||| || |||||||| ||| ||| |||||| ||
Qy         162 actgtagacacaagaaattcaagggatcggttctgtgaaactgtgttccctgaaccagttgtg 221
               || || |||| ||| |||| ||||||| |||||| ||||||||| || ||| ||||||||| ||
Db         189 ATTGCAGGCAGCGAGGAATTCAAGGATCGATTCTGGAAACTGTGTTCCCTGCAAACAGTTGCG 248
Qy         222 ggccaggcaatggagttgtctaaggaatgtggctttcggctatgtggggaggatgcacagtgtg 281
               |||||||| ||| |||||||||||||||| || ||||||||||||| ||||||| |||| |
Db         249 GACCTGGCCATGGAGTTGTCTCAAGGAAATGTGGCTTTCGGCTATNGGGGAGGANTGCACAGTGTG 308
Qy         282 tggcgtgccggctgcacaggttcaacaggtggggcttccaagaatg 330
               ||||||||| ||| |||||||||| |||||||| || |||| ||
Db         309 TGCCCGTGCCAGGCGCGCCGCCACCGGTTCAAGGAAGACTGGGGGTTTCAAGAAGTG 357
```

RESULT 6

```
Db    375 GCGCTGTGCTCGCGTATCTCGCCGCTCTCGGCCACCGTCCTCCTCTGCCTGGCCTCATCCTC
          |||||       |           ||||||||||||||||||||||||||||||||||||||
Qy    629 tgtgtcatctattgtaagagacagttatggaaagagaaaccagctggttctctgcggtca  688
Db    435 TGTGTCATCTCATTGTAAGAGACAGTNTATGAGAGAAGAAAACCCAGCTGTAAGCTGTTCAGCT  494
Qy    689 caggacattcagtacaac  706
          ||||  ||||| ||  ||
Db    495 CATTACAINTCTTATCAC  512
```

LOCUS   AV111112   275 bp   mRNA   EST   29-JUN-1999
DEFINITION   AV111112 Mus musculus C57BL/6J 10-day embryo Mus musculus cDNA
        clone 2600016N17, mRNA sequence.
ACCESSION   AV111112
VERSION   AV111112.1   GI:5265192
KEYWORDS   EST.
SOURCE   house mouse.
   ORGANISM   Mus musculus
       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
       Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 275)
  AUTHORS   Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
       Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T.,
       Hara,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M.,
       Kawai,J., Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H.,
       Owa,C., Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y.,
       Sugahara,Y., Suzuki,H., Suzuki,Y., Tateno,M., Tomaru,Y.,
       Tominaga,N., Watanabe,S., Yagame,M., Yamamura,T., Yokota,T.,
       Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
  TITLE   RIKEN Mouse ESTs
  JOURNAL   Unpublished (1999)
COMMENT   On Jun 5, 1998 this sequence version replaced gi:3187126.
       Contact: Chie Owa
       Genome Science Laboratory
       RIKEN
       3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
       Tel: 81-298-36-9145
       Fax: 81-298-36-9098
       Email: genome-res@rtc.riken.go.jp
       Thermostabilization and thermoactivation of thermolabile enzymes by
       trehalose and its application for the synthesis of full length cDNA
       (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
       Transcriptional sequencing: A method for DNA sequencing using RNA
       polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
       Please visit our web site (http://genome.rtc.riken.go.jp) for
       further details.
FEATURES   Location/Qualifiers
  source   1..275
       /organism="Mus musculus"
       /strain="C57BL/6J"
       /db_xref="taxon:10090"
       /clone="2600016N17"
       /clone_lib="Mus musculus C57BL/6J 10-day embryo"
       /sex="mixed"
       /dev_stage="10-day embryo"
BASE COUNT   77 a   86 c   55 g   57 t
ORIGIN

Query Match   8.4%;   Score 124.4;   DB 59;   Length 275;
Best Local Similarity   71.3%;
Matches 164;   Conservative   0;   Mismatches 66;   Indels   0;   Gaps   0;

```
Qy    309 aggactcggggcttccacagaaatgcaagccctgtcgttcggactgcgagtggtgaaccgcttc  368
          ||  ||| ||||||| || |||| | || || ||||||| ||  ||||||| ||||||||
Db     43 AAGAAAGTGGTTTCCAGAAGTAAAAGCCATTCGGCGAACCTGCGCCTGTGTAGAACCCCTATC  102
Qy    369 agaaggcaaattgttcagccacacagtgtgatgcatctcgggggactgcttgccaggattt  428
          ||  ||| || ||||| ||    |||||||||||||||||||||||||| || |||||||||
Db    323 CTCAGCTCCACGAATATGCCCACAGAGCCTGCTGCCAGTGCCTGAGACTCAGTGCAGA  382
```

RESULT 7

```
Db    103 TGAGGACCCATTCCTCACACACCACTGATGCACTCTGCGGGGGGACTGCCTGCCAAGGATTT  162
          ||||||  || ||| || |     | |  | || || |||||| ||||| ||| |||||||
Qy    429 ataggaagacgaaacttgtcgcgtcttcaagacaatggaagtgtgtcgccttgtgaggaccctc  488
Db    163 ACCAGAAGACACAAACTGGTTGGTTTTCAAGACAINGGAGTGTGTGCCCTGCGAGAACCCAC  222
Qy    489 ctcctccttucgaaccgcaactgtgccagcaaggtcaacctcgtgaaagatc  538
          |||| || || | | |||| |   |||  || || |||| | ||| || |
Db    223 CTCCTCCTTACGAACCACACACAGTGAGTGATGTTGCCAAGTGGCAGCAGAC  272
```

LOCUS   AQ544065   646 bp   DNA   GSS   19-MAY-1999
DEFINITION   RPCI-11-315F10.TV RPCI-11 Homo sapiens genomic clone
       RPCI-11-315F10, genomic survey sequence.
ACCESSION   AQ544065
VERSION   AQ544065.1   GI:4869459
KEYWORDS   GSS.
SOURCE   human.
   ORGANISM   Homo sapiens
       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
       Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 646)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
       Venter,J.C.
  TITLE   Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
       Map Building
  JOURNAL   Unpublished (1997)
COMMENT   On Mar 23, 1999 this sequence version replaced gi:3324949.
       Other_GSSs: RPCI-11-315F10.TJ
       Contact: Shaying Zhao, William Nierman, Mark Adams
       Department of Eukaryotic Genomics
       The Institute for Genomic Research
       9712 Medical Center Dr., Rockville, MD 20850
       Tel: 301 838 0200
       Fax: 301 838 0208
       Email: hbbe@tigr.org
       Clones are derived from the human BAC library RPCI-11. For BAC
       library availability, please contact Pieter de Jong
       (pieter@dejong.med.buffalo.edu). Clones may be purchased from
       BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
       Research Genet cs (info@resgen.com). BAC end search page:
       http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
       Seq primer: T7
       Class: BAC ends.
FEATURES   Location/Qualifiers
  source   1..646
       /organism="Homo sapiens"
       /db_xref="GDB:7620705"
       /db_xref="taxon:9606"
       /clone="RPCI-11-315F10"
       /clone_lib="RPCI-11"
       /sex="Male"
       /cell_type="Lymphocytes"
       /note="vector; pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
       RPCI11 Human Male BAC Library"
BASE COUNT   175 a   135 c   145 g   191 t
ORIGIN

Query Match   8.2%;   Score 122.6;   DB 104;   Length 646;
Best Local Similarity   93.4%;
Matches 128;   Conservative   0;   Mismatches   9;   Indels   0;   Gaps   0;

```
Qy    675 ggtctctgccgtccaccaggacaattcagttacaacgagactgagctgctcgtcgtttgacagac  734
          ||||||| | |||||| || |||||||||||| |||| ||| |||||| ||||||||||||
Db    263 GGTCCTCTGCtGTCACAGGACATTCAGTACAACACGGGCTCTCGAGCCTGTCGGTGTTTTGACAGAC  322
Qy    735 ctcagctccacgaatatgcccacaagagacctgccagtgccccgtgactcagtgaga  794
          ||||||||||||| |||| || |||| ||||| ||||| |||| ||||||||||||
Db    323 CTCAGCTCCACGAATATGCCCACAGAGCCTGCTGCCAGTGCCTGAGACTCAGTGCAGA  382
```

```
RESULT  8

Qy    795 cctgccggggccggtcgcgc 811
          |||||||||| || ||
Db    383 CCTGCGGTAAGTTCAGC 399

Qy    795 cctgcggggccggtcgcgc 811
          |||||||||| || ||
Db    383 CCTGCGGTAAGTTCAGC 399
```

LOCUS   AA495217   404 bp   mRNA   EST   27-JUN-1997
DEFINITION fa04d10.r1 zebrafish ICRZfis Danio rerio cDNA clone 10E17 5', mRNA
            sequence.
ACCESSION   AA495217
VERSION   AA495217.1 GI:2225645
KEYWORDS   EST.
SOURCE   zebrafish.
ORGANISM   Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
            Euteleostomi; Actinopterygii;
            Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
            Cyprinoidea; Cyprinidae; Rasborinae; Danio.
REFERENCE   1 (bases 1 to 404)
AUTHORS   Clark,M., Lehrach,H., Appel,B., Eisen,J., Johnson,S., Marra,M.,
            Eddy,S., Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G.,
            Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Martin,J., Moore,B.,
            Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y.,
            Wylie,T., Waterston,R. and Wilson,R.
TITLE   WashU Zebrafish EST Project
JOURNAL   Unpublished (1997)
COMMENT   On May 9, 1995 this sequence version replaced gi:802250.
            Contact: Steve Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Steve Johnson lab internal ID - P1_332 NOTE -- For this library, the
            CLONE id field represents a position identifier on the original
            cDNA library preparation plate. cDNA Library Preparation: Matthew
            Clark, cDNA Library Arrayed by: Matthew Clark. DNA Sequencing by:
            Washington University Genome Sequencing Center Clone distribution:
            Genome Systems, St. Louis, and Max Planck Institut fuer Molekulare
            Genetik, Berlin Tel +49 30 84 13 1235
            Seq primer: T7 ET from Amersham
            High quality sequence stop: 320.
FEATURES   Location/Qualifiers
     source   1..404
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="10E17"
                     /clone_lib="Zebrafish ICRZfis"
                     /sex="mixed"
                     /tissue_type="pooled 26-somite embryos"
                     /note="Vector: pSPORT1: Site_1: NotI; Site_2: SalI; 1st
                     strand cDNA was primed with a Not I - oligo(dT)15 primer
                     [5'pGACTAGTTCTAGATCGCGAGCGGCCGCCCCTTTTTTTTTTTTTTT3'], on
                     mRNA from pooled 26 somite zebrafish embryos;
                     double-stranded cDNA was ligated to Sal I adaptors (BRL),
                     digested with Not I and cloned into the Not I and Sal I
                     sites of the pSPORT1 vector (BRL). Library was constructed
                     by Matthew Clark (Lehrach lab; ICRF, London and Max
                     Planck Institut fuer Molekulare Genetik, Berlin) and was
                     not biochemically normalised. 70,000 clones from this
                     library were arrayed on high density filters and
                     subsequently screened by oligonucleotide hybridization
                     fingerprinting to identify unique or minimally redundant
                     clones for more intensive analysis."
BASE COUNT   80 a   98 c   125 g   101 t
ORIGIN

Query Match   7.8%; Score 116; DB 34; Length 404;
Best Local Similarity 67.0%; Pred. No. 4.2e-20;
Matches 197; Conservative 0; Mismatches 90; Indels 7; Gaps 2;

```
RESULT  9

Qy    560 ctcatcctctgtcatctatctgttatctgcagacggctctggccacgtct 609
          || |||  ||||||| | |||||||||||||||||||||||||||| |||
Db      1 CGAGACAATGGCTCTGGCCTGCCTGGCTGTGATCTTACTGTAAGACACAGCTGCTGGGACAGAGAAACC--AGTATCC 117

Qy    620 ctggcaacaggcgctgggtcgtttatctgcagagggcctcggccacgtct 679
          |||||  ||||| ||||||| | || |||| |||||| || ||||||||||
Db     61 TTCATCCCCTGTGTGGATCTTGTGATCTTAACTGTAAGAGACACAGCTGCTGGGACAGAGAAACC 117

Qy    680 ctgcggttcaacaggacattcagttacaacgagactgagctgtcgtcgttttgacagacctcag 739
          ||||||||||||||| |||||| |||| |||||||||||||||||  ||||||||||||||
Db    118 ATGAGGGGTCATGAGGGCCCGTTCCTGGCTCGGCAGTTGTCTTCGTTGTTGGACCGGAGAGTG 177

Qy    740 ctccacgaatatgcccacaggagctgctgccagtgccgccgtgactcagtgagacctgc 799
          || ||||||||  |  | | |||||||||| ||||||||| ||||| |||| |  | ||
Db    178 CT-----GGAGTTGTCCAGGAGCACCCTGCTCACTGACCATCATAGGATCAGGACGACAGTGT 233

Qy    800 gggccggtgcgcttgctcccatccatgtgctgtgaggaggcctgcagcgcccac 853
          |||||||||||||| | ||||||||||| ||| || ||| || ||  ||
Db    234 GGTGGCGTTCAGCTGGGTCGTCGTCGTTGTTGGAGATGTCTGGAGTCAAAAC 287
```

LOCUS   AQ514075/c   450 bp   DNA   GSS   05-MAY-1999
DEFINITION HS_5184_A1_G01_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=760 Col=1 Row=M, genomic survey sequence.
ACCESSION   AQ514075
VERSION   AQ514075.1 GI:4746366
KEYWORDS   GSS.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 450)
AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE   Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE   99380589
COMMENT   Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
FEATURES   Location/Qualifiers
     source   1..450
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate-760 Col-1 Row=M"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Genomic sequence of BAC ends"
            Seq primer: SP6   row: M   column: 1
            Plate: 760   row: M   column: 1
            Class: BAC ends
            High quality sequence stop: 450.
BASE COUNT   156 a   93 c   62 g   137 t   2 others
ORIGIN

Query Match   7.5%; Score 112.2; DB 104; Length 450;
Best Local Similarity 93.6%; Pred. No. 4.3e-19;
Matches 117; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

RESULT 10

```
Qy   128 tatttgtcatgtaaagtgacttgtgaatccagcaggegagactggagacgcaagagaattcaggat 187
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   327 TATTCAGTCATGTAAAGTGACTTGTGATACAGGAGACTGGAGACACAAGAGAATTCAGGGAT 268

Qy   188 cggttctgggaaactgtgttccctgcaaccagtgggcaggaggggagttgtctaaggaa 247
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   267 CGGTTCTGGGAAAATGTGTTCCTGCAACCAGTGTGGGCCCAGGCCATGGAGTTGTCTAAGGTA 208

Qy   248 tgtgg 252
         |||||
Db   207 TATTG 203
```

LOCUS   AI551729   381 bp   mRNA   EST   23-MAR-1999
DEFINITION vf11e02.y1 Knowles Solter mouse blastocyst B3 Mus musculus cDNA
            clone IMAGE:835418 5', mRNA sequence.
ACCESSION   AI551729
VERSION   AI551729.1   GI:4484092
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 381)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE   The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
  COMMENT   On Mar 10, 1998 this sequence version replaced gi:2948459.
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:495634
            This read is a RESEQUENCE of a previously sequenced mouse clone
            This read has been verified (found to hit its original self in the
            correct orientation)
            Seq primer: -40RP.
FEATURES   Location/Qualifiers
     source   1..381
                     /organism="Mus musculus"
                     /strain="B6D2 F1/J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:835418"
                     /clone_lib="Knowles Solter mouse blastocyst B3"
                     /tissue_type="blastocyst"
                     /dev_stage="embryo (pre-implantation)"
                     /note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally from mRNA prepared
                     from 800 blastocysts. Primer: SalI(dT):
                     5'-CGGTCGACGCTGCGACCGTTTTTTTTTTTTTTT-3'. cDNAs were
                     cloned into the NotI/SalI sites of a pSPORT vector (life
                     technologies). Two different size selections: B1 (larger
                     inserts) and B3."
BASE COUNT   105 a   95 c   104 g   77 t
ORIGIN

Query Match   5.3%;   Score 78.6; DB 48;   Length 381;
Best Local Similarity 55.8%;   Pred. No. 3e-10;
Matches 150; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
```

RESULT 11

```
Qy  1042 tcctgaactcactggaagaagacattcattctctccaatccagaacttgaaagctcaacgtc 1101
         |||||||| || |||||||||| || |||||||| ||||||  ||||| |||| |||| ||
Db     1 TCCTGAACTCACTGGAGAAGATACCAAATTCCCTCCAATCCCGAAAACGAAAAGCCCAGCATC 60

Qy  1102 tttggattcaaatagcagtcaaagatttggttggttgggg cgttccagtcctccattc 1161
         ||||||||| |||||||| ||||| |||||||||||  || || |  ||||||| ||| |
Db    61 TCTGGATTCCAGTGGCGGCGGGGCCAGGATCTGGCTGGGGACACAGGTGCTCTTAGAGTCTTCGGAA 120

Qy  1162 tgaaaacttacagcagctcactgatttatctagatataaacacacactggtagaatcagc 1221
         ||| |||||| |||||| ||||||||| || |||||||| || |||||| ||  |||||
Db   121 TGTTTCAGAATCTACTGACTTCACCTAGACACATGGTGCACACTGGTACAGTCTGGGACACAGAC 180

Qy  1222 atcaactcaggagtgcactaactatgaagagccagctagaatcaggaggagtggcgctgtcat 1281
         ||||||||| ||| | |||||||| |||  || |||  |||| ||||| || |||  ||| |
Db   181 GCTRGCTCAGGATGCTCAAGAAGACTCCAAGCAAGACGGCCTGGGAAGACAGGAGAAAACCT 240

Qy  1282 ccaccccagccacctcaagacgtccctccagg 1310
         |||||||||| |||||||| ||  || || |||
Db   241 GAATCTAGCCATGCCCACACCCTTCCAGG 269
```

LOCUS   AL120773   485 bp   mRNA   EST   27-SEP-1999
DEFINITION DKFZp763c192_r1 762 (synonym: hmel2) Homo sapiens cDNA clone
            DKFZp762c192 5', mRNA sequence.
ACCESSION   AL120773
VERSION   AL120773.1   GI:5926774
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 485)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE   EST (Koehrer, et al.)
  JOURNAL   Unpublished (1999)
  COMMENT   On Mar 16, 1998 this sequence version replaced gi:2961733.
            Contact: Koehrer K
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann; Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMFZ within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rpd.de.
FEATURES   Location/Qualifiers
     source   1..485
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp762c192"
                     /clone_lib="762 (synonym: hmel2)"
                     /tissue_type="melanoma (MeWo cell line)"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSportl; Site_1: NotI; Site_2: SalI"
BASE COUNT   150 a   103 c   104 g   128 t
ORIGIN

Query Match   4.2%;   Score 61.8; DB 64;   Length 485;
Best Local Similarity 90.4%;   Pred. No. 8.7e-06;
Matches 66; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```
Qy   15 ccaggtggctggaagaactctccaacaataatacacatttgataagaaagatggcttaa 74
         || |||||| | ||||||| ||||||||| ||| || ||||| ||||| |||| |
Db   413 CTAAGTTGCTGAACGGAACTCTCCAACAATAAAATACATTGATAAGAAAAGAATGGCTTTAA 472

Qy   75 aagtgctactaga 87
```

```
Db     473 AAGTGCTACTAGA 485
           |||||||||||||
```

RESULT 12
AQ173919/c

```
LOCUS       AQ173919        405 bp    DNA            GSS        17-OCT-1998
DEFINITION  HS_3204_A2_C10_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3204 Col=20 Row=E, genomic survey
            sequence.
ACCESSION   AQ173919
VERSION     AQ173919.1  GI:3571286
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 405)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., and
            Hood,L.
TITLE       Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE     99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3204 row: E column: 20
            Class: BAC ends
            High quality sequence stop: 405.
FEATURES             Location/Qualifiers
     source          1..405
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=3204 Col=20 Row=E"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-coli DH10B"
BASE COUNT    109 a    98 c    112 g    85 t    1 others
ORIGIN

Query Match     4.0%;  Score 59;  DB 99;  Length 405;
Best Local Similarity   92.5%;  Pred. No. 4.6e-05;
Matches   62;  Conservative  0;  Mismatches  5;  Indels  0;  Gaps  0;

Qy    454 tcaagacatggagtcgtgtcctgtgtgggagaccgcactgtgc 513
          ||||||||||||||||||||||||||||| |||||||||||||
Db    405 TCAAGACATGGAGTCGTGTCCTGTGGAGACCCTCCCTCCTTCTTACGAACCGCACACTGTGA 346

Qy    514 cagcaag 520
          ||| |||
Db    345 GTGAACG 339
```

RESULT 13
CN500LPM/c

```
LOCUS       CN500LPM        1101 bp    DNA            GSS        03-JUN-1999
DEFINITION  CN500LPM Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR32201 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL068516
VERSION     AL068516.1  GI:4958747
KEYWORDS    GSS.
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp; the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR32J01"
                     /note="end : TET3"
BASE COUNT    350 a    161 c    165 g    254 t    171 others
ORIGIN

Query Match     2.9%;  Score 43.4;  DB 79;  Length 1101;
Best Local Similarity   61.4%;  Pred. No. 0.74;
Matches   51;  Conservative  10;  Mismatches  22;  Indels  0;  Gaps  0;

Qy   1407 tttttttttgcatcttaataattcttatatgttgtaggtatgtttaaaataaatt 1466
          ||||||||| |:|||| ||:|||||| |||||||||  |||||||| ||||:|||
Db    327 TTTTTTTTTAGMWTTTTTTATBATTTATKATTTGTTTTACASCWTTTTTACTCBATTAWTC 268

Qy   1467 caagtatttttttaaaaaacttt 1489
          | :||:| | ||||| || |||
Db    267 TTACTAWTKTAWTGAATAATTTT 245
```

RESULT 14
AA759377/c

```
LOCUS       AA759377        443 bp    mRNA           EST        29-DEC-1998
DEFINITION  ah54al0.s1 Soares_testis_NHT Homo sapiens cDNA clone 1309434 3',
            mRNA sequence.
ACCESSION   AA759377
VERSION     AA759377.1  GI:2804852
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 443)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2151262.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
```

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 587   Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 424.

FEATURES   Location/Qualifiers
     source   1. .443
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="1309434"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech
                     laboratories, Inc., and primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCGGAAGTGGAGGAGGCGGCGCCCCCAAATTTTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT   119 a   107 c   117 g   98 t   2 others
ORIGIN

Query Match   2.9%;   Score 42.6; DB 38; Length 443;
Best Local Similarity   65.9%;   Pred. No. 0.99;
Matches   60; Conservative   0; Mismatches 31; Indels   0; Gaps   0;

Qy   537 tcggtccacggcctccagcccacgggggacacggcggctggcgtcggcggttatctgcagcggctc 596
           ||||| ||||||||| ||| ||||||| ||| |||||||||| ||||||||||||||||||||||||
Db   416 TCGGGGCCACGGGCCGGGCATCANNCCCGTCACCCGGCTCTCTACGGCTTCATCTTCTACCTGC 357

Qy   597 tggccaccgtcctgctggccctgctcatcct 627
           |||||||||||||||||||||||||||||||
Db   356 TCGGCCTCCTCCTCGCTCTCCCTGCTCCTCAT 326

LOCUS   CNS009YZ   806 bp   DNA   GSS   03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
            BACR20021 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL054514
VERSION   AL054514.1   GI:4935687
KEYWORDS   GSS.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 806)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the library
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila/drosophila_bac.htm.

FEATURES   Location/Qualifiers
     source   1. .806
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR20021"
                     /note="end : T7"

BASE COUNT   288 a   86 c   116 g   160 t   156 others
ORIGIN

Query Match   2.8%;   Score 42.4; DB 79; Length 806;
Best Local Similarity   38.6%;   Pred. No. 1.3;
Matches   78; Conservative 34; Mismatches 90; Indels   0; Gaps   0;

Qy   1288 agccactcagacgtccctccaggtaaggcagcgactgggttccctgtgaacacagcactg 1347
           ||| | | || | || | | | | |||| |||||| ||| || | | |||||
Db   606 HSMVACTMRKMWVTCMTGCGMCRCRCKCBASYGTCYCYCYATACATMTCTAYCATATY 547

Qy   1348 acttacagtagataagaaactctgttcccagccataagattggggggaacctgaatgaagtt 1407
           || | |||| ||: :||| |||| :| ||:||||| || | :|| | | |||| | |||
Db   546 ACAYIMWTATATATMMAMYIYTCTMCCCWAYVTCWCATCWMATTTTTTTTTTTTTTTTTTT 487

Qy   1408 ttttttttg:atctttaataaattcttatatgttgtagagtatgtttaaaataaattc 1467
           ||||||||| |||| |||| | ||| ||| |||| | || | ||||||||||||
Db   486 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 427

Qy   1468 aagtattttttaaaaacttt 1489
           || || |||| |||: : : :
Db   426 TTTTTTTTTTTTYTTTTAAAWTTKK 405

Search completed: April 23, 2000, 01:49:00
Job time: 1275 sec

OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 23, 2000, 13:40:26 ; Search time 28.62 seconds
6228.783 Million cell updates/sec

Title: US-09-490-187-1
Perfect score: 1489
Sequence: 1 ggacctgcagcctcccaggt...........gtatttttttaaaaacttt 1489
Scoring table: IDENTITY_NUC
Scoring table: GapOp 10.0 , GapExt 1.0
Searched: 214294 seqs, 5986157 residues

Total number of hits satisfying chosen parameters: 428588

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :   Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/1na/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/1na/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/1na/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/1na/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/1na/6_COMB.seq:*
6: /cgn2_6/ptodata/1/1na/PCTUS9_COMB.seq:*
7: /cgn2_6/ptodata/1/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39.6	2.7	4673	1	US-07-638-431-1	Sequence 1, Appli
2	39.6	2.7	4673	6	PCT-US92-00018-1	Sequence 1, Appli
3	36.6	2.5	761	2	US-08-477-87B-83	Sequence 83, Appl
4	36.6	2.5	761	2	US-08-472-281A-83	Sequence 83, Appl
5	36.6	2.5	761	4	US-08-477-989B-83	Sequence 83, Appl
6	36.6	2.5	761	4	US-08-477-87B-90	Sequence 90, Appl
7	36.6	2.5	807	2	US-08-472-281A-90	Sequence 90, Appl
8	36.6	2.5	807	4	US-08-477-989B-90	Sequence 90, Appl
9	36.4	2.4	807	4	US-08-722-001-7	Sequence 7, Appl
10	36.4	2.4	1601	2	US-08-722-001-26	Sequence 26, Appl
11	36.4	2.4	1987	2	US-08-722-001-27	Sequence 27, Appl
12	36.4	2.4	1997	2	US-08-722-001-11	Sequence 11, Appl
13	36.2	2.4	2004	2	US-08-424-424B-1	Sequence 1, Appl
14	36.2	2.4	2485	6	PCT-US94-05363A-1	Sequence 1, Appl
15	35.4	2.4	2486	6	US-08-705-937-10	Sequence 10, Appl
16	34.8	2.3	1183	1	US-08-334-698-5	Sequence 5, Appl
17	34.8	2.3	1639	1	US-08-228-932-5	Sequence 5, Appl
18	34.8	2.3	1639	1	US-08-466-939-5	Sequence 5, Appl
19	34.8	2.3	1639	3	US-08-406-855A-5	Sequence 5, Appl
20	34.8	2.3	1639	5	US-08-722-190-5	Sequence 5, Appl
21	34.8	2.3	1639	6	PCT-US95-04203-5	Sequence 5, Appl
22	34.6	2.3	1542	4	US-09-122-230-8	Sequence 8, Appl
23	34.6	2.3	1909	4	US-09-122-230-6	Sequence 6, Appl
24	34.6	2.3	1909	4	US-08-233-463-14	Sequence 14, Appl
25	34.2	2.3	7718	2	US-08-487-826B-13	Sequence 13, Appl
26	33.8	2.3	19124	4	US-08-582-257-20	Sequence 20, Appl
27	33.8	2.3	568	3	US-08-582-298-20	Sequence 20, Appl
28	33.8	2.3	1493	1	US-08-340-820-24	Sequence 24, Appl
29	33.8	2.3	1493	1	US-08-593-535-24	Sequence 24, Appl
30	33.8	2.3	1740	2	US-08-362-512A-3	Sequence 3, Appl
31	33.8	2.3	2861	2	US-08-299-953-1	Sequence 1, Appl
32	33.8	2.3	2861	2	US-08-459-415-1	Sequence 1, Appl
33	33.8	2.3	2861	6	PCT-US95-11211-1	Sequence 1, Appl
34	33.8	2.3	3881	2	US-08-259-953-2	Sequence 2, Appl
35	33.8	2.3	3881	2	US-08-459-415-2	Sequence 2, Appl
36	33.8	2.3	3881	6	PCT-US95-11211-2	Sequence 2, Appl
37	33.4	2.3	2762	1	US-08-198-446B-12	Sequence 12, Appl
38	33.2	2.2	2762	3	US-08-870-693-12	Sequence 12, Appl
39	33.2	2.2	1593	1	US-08-307-499-50	Sequence 50, Appl
40	33.2	2.2	8174	1	US-07-914-281-5	Sequence 5, Appl
41	33.2	2.2	8174	1	US-08-393-246-5	Sequence 5, Appl
42	33.2	2.2	8174	4	US-08-525-058A-5	Sequence 5, Appl
43	33.2	2.2	8174	5	US-08-696-731-5	Sequence 5, Appl
44	33.2	2.2	8174	6	PCT-US91-00899-3	Sequence 3, Appl
45	33.2	2.2	14176	1	US-08-307-499-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-638-431-1
Sequence 1, Application US/07638431
Patent No. 5198535
GENERAL INFORMATION:
APPLICANT: Hoffman, Stephen L.
APPLICANT: Charoenvit, Yupin
APPLICANT: Hedstrom, Richard
APPLICANT: Khusmith, Srisin
APPLICANT: Rogers IV, William O.
TITLE OF INVENTION: Protective malaria sporozoite surface protein
TITLE OF INVENTION: Immunogen and gene
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: A. David Speyack
STREET: NMRDC Building 1 T-12 National Naval
STREET: Medical Center
CITY: Bethesda
STATE: MD
COUNTRY: USA
ZIP: 20814-5044
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/638,431
FILING DATE: 19910110
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Speyack, Avrom D.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 295-6759
TELEFAX: (301) 295-4033
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4673 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N
ANTI-SENSE: N
ORIGINAL SOURCE:
ORGANISM: Plasmodium yoelii
STRAIN: 17X(NL)
DEVELOPMENTAL STAGE: erythrocytic stage
TISSUE TYPE: Blood
CELL TYPE: erythrocytic stage

; IMMEDIATE SOURCE:
;   LIBRARY: Py-lambdagt11-2-7 kb genomic expression
;   CLONE: Py10.1111
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 718..3195
; OTHER INFORMATION:
US-07-638-431-1

Qy 1379 ataagatttgggggaacctggatggagtgttttttttttgcatcttaataaattccttatat 1438
         |||||||||||||||||||||||||| |||||||||||||| || ||||||||| || ||
Db 4033 ATATTATTTGTGTAAAAAATTAAGATGATATATTTTAGCATATTTTGACAAAATTGTCAAAT 4092

Qy 1439 gttgtaggagtatgttttaaaataaattcaagtatttttttaaaaaacctt 1488
         |||||| |||||||||||||||||||||||||||||||||||||||||
Db 4093 TCGTTGTATTCTGTTTTCGAAAAATATGAGCATTTTTTTTTTAATATTATT 4142

Query Match   2.7%; Score 39.6; DB 1; Length 4673;
Best Local Similarity 60.0%; Pred. No. 0.085;
Matches 66; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

RESULT 2
PCT-US92-00018-1
; Sequence 1, Application PC/US9200018
; GENERAL INFORMATION:
;   APPLICANT: Hoffman, Stephen L.
;   APPLICANT: Charoenvit, Yupin
;   APPLICANT: Hedstrom, Richard
;   APPLICANT: Khusmith, Srisin
;   APPLICANT: Rogers IV, William O.
;   TITLE OF INVENTION: Protective malaria sporozoite surface protein
;   TITLE OF INVENTION: Immunogen and gene encoding
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: A. David Spevack
;   STREET: NMRDC Building 1 T-12 National Naval
;   STREET: Medical Center
;   CITY: Bethesda
;   STATE: MD
;   ZIP: 20814-5044
;   COUNTRY: USA
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.24
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US92/00018
;   FILING DATE: 19920103
;   CLASSIFICATION: 424
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Spevack, Avram D.
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (301) 295-6759
;   TELEFAX: (301) 295-4033
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4673 base pairs
;   TYPE: NUCLEIC ACID
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   HYPOTHETICAL: N
;   ANTI-SENSE: N
;   ORIGINAL SOURCE:
;   ORGANISM: Plasmodium yoelii
;   STRAIN: 17X(NL)
;   DEVELOPMENTAL STAGE: erythrocytic stage
;   TISSUE TYPE: Blood
;   CELL TYPE: erythrocytic stage
;   IMMEDIATE SOURCE:

; IMMEDIATE SOURCE:
;   LIBRARY: Py-lambdagt11-2-7 kb genomic expression
;   CLONE: Py10.1111
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 718..3195
; OTHER INFORMATION:
PCT-US92-00018-1

Qy 1379 ataagatttgggggaacctggatggagtgttttttttttgcatcttaataaattccttatat 1438
         |||||||||||||||||||||||||| |||||||||||||| || ||||||||| || ||
Db 4033 ATATTATTTGTGTAAAAAATTAAGATGATATATTTTAGCATATTTTGACAAAATTGTCAAAT 4092

Qy 1439 gttgtaggagtatgttttaaaataaattcaagtatttttttaaaaaacctt 1488
         |||||| |||||||||||||||||||||||||||||||||||||||||
Db 4093 TCGTTGTATTCTGTTTTCGAAAAATATGAGCATTTTTTTTTTAATATTATT 4142

Query Match   2.7%; Score 39.6; DB 6; Length 4673;
Best Local Similarity 60.0%; Pred. No. 0.085;
Matches 66; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

RESULT 3
US-08-477-877B-83/c
; Sequence 83, Application US/08477877B
; Patent No. 5730979
; GENERAL INFORMATION:
;   APPLICANT: Bazin, Herv
;   APPLICANT: Latinne, Dominique
;   TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Act
;   NUMBER OF SEQUENCES: 96
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
;   ADDRESSEE: Cecchi, Stewart & Olstein
;   STREET: 6 Becker Farm Road
;   CITY: Roseland
;   STATE: New Jersey
;   COUNTRY: U.S.A.
;   ZIP: 07068
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch diskette
;   COMPUTER: IBM PS/2
;   OPERATING SYSTEM: MS-DOS
;   SOFTWARE: WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/477,877B
;   FILING DATE: 07-JUN-1995
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/407,009
;   FILING DATE: 29-MAR-1995
;   APPLICATION NUMBER: 08/119,032
;   FILING DATE: 09-SEP-1993
;   APPLICATION NUMBER: 08/027,008
;   FILING DATE: 05-MAR-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Olstein, Elliot M.
;   REGISTRATION NUMBER: 24,025
;   REFERENCE/DOCKET NUMBER: 61750-146
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 201-994-1700
;   TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 83:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 761 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: oligonucleotide
;   FEATURE:
;   NAME/KEY: US-08-477-877B-83
;   MOLECULE TYPE: Nucleotide sequence encoding LO-CD2a VL chain.

Query Match
Best Local Similarity 54.0%;   Score 36.6;   DB 2;   Length 761;
Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
Pred. No. 0.21;

```
Qy  1348  acttacacagtagatcagaacctctgttcccagcataagaattgggggaacctggatcgagtt  1407
              ||||||  |||||||||||||  |||  ||| ||     || |  |||
Db   410  AATTATAATAGATCAGAACAAAGTTATGTCAAAATTTGCAACCAATCTTTGTGAATTAC  351

Qy  1408  tttttttgcatcttaataattcttatatgttgtagagtatgttttaaataaattc  1467
              ||| |||||    |||||||||||||||| || |   |||||||| ||
Db   350  TTTTTGTTACTCATTTCCTGTGAGATCCCGGTGTGTTTTGTGATTTTTAAAAAATATACC  291

Qy  1468  aagtattttttaaaaac  1486
              ||||||| |||||
Db   290  AAATGCTTTTTCACAAAAC  272
```

RESULT 4
US-472-281A-83/c
; Sequence 83, Application US/08472281A
; Patent No. 5817311
; GENERAL INFORMATION:
;   APPLICANT: Bazin, Herv
;   APPLICANT: Latinne, Dominique
;   TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activa
;   NUMBER OF SEQUENCES: 96
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
;   ADDRESSEE: Cecchi, Stewart & Olstein
;   STREET: 6 Becker Farm Road
;   CITY: Roseland
;   STATE: New Jersey
;   COUNTRY: U.S.A.
;   ZIP: 07068
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch diskette
;   COMPUTER: IBM PS/2
;   OPERATING SYSTEM: MS-DOS
;   SOFTWARE: WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/472,281A
;   FILING DATE: 07-JUN-1995
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/407,009
;   FILING DATE: 29-MAR-1995
;   APPLICATION NUMBER: 08/119,032
;   FILING DATE: 09-SEP-1993
;   APPLICATION NUMBER: 08/027,008
;   FILING DATE: 05-MAR-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Olstein, Elliot M.
;   REGISTRATION NUMBER: 24,025
;   REFERENCE/DOCKET NUMBER: 61750-142
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 201-994-1700
;   TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 83:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 761 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: oligonucleotide
;   FEATURE:
;   NAME/KEY: Nucleotide sequence encoding LO-CD2a VL chain.
US-08-472-281A-83

Query Match
Best Local Similarity 54.0%;   Score 36.6;   DB 2;   Length 761;
Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
Pred. No. 0.21;

```
Qy  1348  acttacacagtagatcagaacctctgttcccagcataagaattgggggaacctggatcgagtt  1407
              ||||||  |||||||||||||  |||  ||| ||     || |  |||
Db   410  AATTATAATAGATCAGAACAAAGTTATGTCAAAATTTGCAACCAATCTTTGTGAATTAC  351

Qy  1408  tttttttgcatcttaataattcttatatgttgtagagtatgttttaaataaattc  1467
              ||| |||||    |||||||||||||||| || |   |||||||| ||
Db   350  TTTTTGTTACTCATTTCCTGTGAGATCCCGGTGTGTTTTGTGATTTTTAAAAAATATACC  291

Qy  1468  aagtattttttaaaaac  1486
              ||||||| |||||
Db   290  AAATGCTTTTTCACAAAAC  272
```

RESULT 5
US-08-477-989B-83/c
; Sequence 83, Application US/08477989B
; Patent No. 5951983
; GENERAL INFORMATION:
;   APPLICANT: Bazin, Herv
;   APPLICANT: Latinne, Dominique
;   APPLICANT: Kaplan, Ruth
;   APPLICANT: Kieber-Emmons, Thomas
;   APPLICANT: Postema, Christina E.
;   APPLICANT: White-Scharf, Mary
;   TITLE OF INVENTION: LO-CD2a Antibody and Uses
;   TITLE OF INVENTION: Thereof for Inhibiting
;   TITLE OF INVENTION: T-Cell Activation and
;   TITLE OF INVENTION: Proliferation
;   NUMBER OF SEQUENCES: 96
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
;   ADDRESSEE: Cecchi, Stewart & Olstein
;   STREET: 6 Becker Farm Road
;   CITY: Roseland
;   STATE: New Jersey
;   COUNTRY: U.S.A.
;   ZIP: 07068
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch diskette
;   COMPUTER: IBM PS/2
;   OPERATING SYSTEM: MS-DOS
;   SOFTWARE: WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/477,989B
;   FILING DATE: 07-JUN-1995
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/407,009
;   FILING DATE: 29-MAR-1995
;   APPLICATION NUMBER: 08/119,032
;   FILING DATE: 09-SEP-1993
;   APPLICATION NUMBER: 08/027,008
;   FILING DATE: 05-MAR-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Olstein, Elliot M.
;   REGISTRATION NUMBER: 24,025
;   REFERENCE/DOCKET NUMBER: 61750-147
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 201-994-1700
;   TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 83:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 761 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: oligonucleotide
;   FEATURE:
;   NAME/KEY: Nucleotide sequence encoding LO-CD2a
;   NAME/KEY: VL chain.
US-08-477-989B-83

Query Match   2.5%;   Score 36.6; DB 4;   Length 761;
Best Local Similarity   54.0%; Pred. No. 0.21;
Matches   75; Conservative   0; Mismatches   64; Indels   0; Gaps   0;

Qy 1348 acttacagtagatcagaactctgttcccagcataagatttggggggaacctggatgagtt 1407
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   410 AATTATAATAGATCAGAACAAAGTATATGTGCAACCAAACTTTGTGAATTAC 351

Qy 1408 tttttttgcatcttaataatttcttatatgttgtagagtatgtttttaaaataaatttc 1467
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   350 TTTTTGTTACTCATTTCCTGTGAGATCCCGGTGTGTTTTGTGATTTTTAAAAAATATACC 291

Qy 1468 aagtatttttttaaaaac 1486
          ||||||||||||||||||
Db   290 AAATGCTTTTCACAAAAC 272

RESULT 6
US-08-477-877B-90/c
; Sequence 90, Application US/08477877B
; Patent No. 5730979
; GENERAL INFORMATION:
;   APPLICANT: Bazin, Herv
;   APPLICANT: Latinne, Dominique
;   TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activa
;   NUMBER OF SEQUENCES: 96
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
;   ADDRESSEE: Cecchi, Stewart & Olstein
;   STREET: 6 Becker Farm Road
;   CITY: Roseland
;   STATE: New Jersey
;   COUNTRY: U.S.A.
;   ZIP: 07068
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch diskette
;   COMPUTER: IBM PS/2
;   OPERATING SYSTEM: MS-DOS
;   SOFTWARE: WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/477,877B
;   FILING DATE: 07-JUN-1995
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/407,009
;   FILING DATE: 29-MAR-1995
;   APPLICATION NUMBER: 08/119,032
;   FILING DATE: 09-SEP-1993
;   APPLICATION NUMBER: 08/027,008
;   FILING DATE: 05-MAR-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Olstein, Elliot M.
;   REGISTRATION NUMBER: 24,025
;   REFERENCE/DOCKET NUMBER: 61750-146
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 201-994-1700
;   TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 90:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 807 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: polynucleotide
;   FEATURE:
;   NAME/KEY: Nucleotide sequence encoding humanized LO-CD2a light chain variable re

Query Match   2.5%;   Score 36.6; DB 2;   Length 807;
Best Local Similarity   54.0%; Pred. No. 0.22;
Matches   75; Conservative   0; Mismatches   64; Indels   0; Gaps   0;

Qy 1348

Qy 1348 acttacagtagatcagaactctgttcccagcataagatttggggggaacctggatgagtt 1407
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   417 AATTATAATAGATCAGAACAAAGTATATGTGCAACCAAACTTTGTGAATTAC 358

Qy 1408 tttttttgcatcttaataatttcttatatgttgtagagtatgtttttaaaataaatttc 1467
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   357 TTTTTGTTACTCATTTCCTGTGAGATCCCGGTGTGTTTTGTGATTTTTAAAAAATATACC 298

Qy 1468 aagtatttttttaaaaac 1486
          ||||||||||||||||||
Db   297 AAATGCTTTTCACAAAAC 279

RESULT 7
US-08-472-281A-90/c
; Sequence 90, Application US/08472281A
; Patent No. 5817311
; GENERAL INFORMATION:
;   APPLICANT: Bazin, Herv
;   APPLICANT: Latinne, Dominique
;   TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Act
;   NUMBER OF SEQUENCES: 96
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
;   ADDRESSEE: Cecchi, Stewart & Olstein
;   STREET: 6 Becker Farm Road
;   CITY: Roseland
;   STATE: New Jersey
;   COUNTRY: U.S.A.
;   ZIP: 07068
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch diskette
;   COMPUTER: IBM PS/2
;   OPERATING SYSTEM: MS-DOS
;   SOFTWARE: WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/472,281A
;   FILING DATE: 07-JUN-1995
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/407,009
;   FILING DATE: 29-MAR-1995
;   APPLICATION NUMBER: 08/119,032
;   FILING DATE: 09-SEP-1993
;   APPLICATION NUMBER: 08/027,008
;   FILING DATE: 05-MAR-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Olstein, Elliot M.
;   REGISTRATION NUMBER: 24,025
;   REFERENCE/DOCKET NUMBER: 61750-142
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 201-994-1700
;   TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 90:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 807 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: polynucleotide
;   FEATURE:
;   NAME/KEY: Nucleotide sequence encoding humanized LO-CD2a light chain variable
US-08-472-281A-90 .;

Query Match   2.5%;   Score 36.6; DB 2;   Length 807;
Best Local Similarity   54.0%; Pred. No. 0.22;
Matches   75; Conservative   0; Mismatches   64; Indels   0; Gaps   0;

Qy 1348 acttacagtagatcagaactctgttcccagcataagatttggggggaacctggatgagtt 1407
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   417 AATTATAATAGATCAGAACAAAGTATGTGCAACCAAATCTTTGTGAATTAC 358

```
Qy   1408 tttttttgcatctttaataattcttatatgttgtagagtagtgtttttaaattaaattc 1467
          ||||| |||  ||| |  ||| | ||  ||||||||||||||||||||||||| || | |
Db    357 TTTTTGTTACTCATTTCCTGTGAGATCCCGGTGTGTTTTGTGATTTTTAAAAAATATACC 298

Qy   1468 aagtattttttaaaaaac 1486
          |||||||||||||||
Db    297 AAAATGCTTTTTCACAAAAC 279
```

```
RESULT   8
US-08-477-989B-90/c
; Sequence 90, Application US/08477989B
; Patent No. 5951983
; GENERAL INFORMATION:
;   APPLICANT: Bazin, Herv
;   APPLICANT: Latinne, Dominique
;   APPLICANT: Kaplan, Ruth
;   APPLICANT: Kieber-Emmons, Thomas
;   APPLICANT: Postema, Christina E.
;   APPLICANT: White-Scharf, Mary
;   TITLE OF INVENTION: Lo-CD2a Antibody and Uses
;   TITLE OF INVENTION: Thereof for Inhibiting
;   TITLE OF INVENTION: T-Cell Activation and
;   TITLE OF INVENTION: Proliferation
;   NUMBER OF SEQUENCES: 96
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
-     ADDRESSEE: Cecchi, Stewart & Olstein
;     STREET: 6 Becker Farm Road
;     CITY: Roseland
;     STATE: New Jersey
;     COUNTRY: U.S.A.
;     ZIP: 07068
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: 3.5 inch diskette
;     COMPUTER: IBM PS/2
;     OPERATING SYSTEM: MS-DOS
;     SOFTWARE: Wordperfect 5.1
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/477,989B
;     FILING DATE: 07-JUN-1995
;     CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/407,009
;     FILING DATE: 29-MAR-1995
;     APPLICATION NUMBER: 08/119,032
;     FILING DATE: 09-SEP-1993
;     APPLICATION NUMBER: 08/027,008
;     FILING DATE: 05-MAR-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Olstein, Elliot M.
;     REGISTRATION NUMBER: 24,025
;     REFERENCE/DOCKET NUMBER: 61750-147
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 201-994-1700
;     TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 90:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 807 bases
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: polynucleotide
;   FEATURE:
;     NAME/KEY: Nucleotide sequence encoding
;     NAME/KEY: humanized LO-CD2a light chain variable region.
US-08-477-989B-90

  Query Match    2.5%;   Score 36.6;  DB 4;  Length 807;
  Best Local Similarity   54.0%;   Pred. No. 0.22;
  Matches   75;  Conservative   0;  Mismatches   64;  Indels   0;  Gaps   0;
```

```
Qy   1408 acttacagtagatcagaactctgttcccagrgatagaggttcggggggaacctggatgagtt 1407
          ||||| ||||  ||| |  ||||  ||| | ||  |||||| ||||||| ||  | | ||
Db    417 AATTATAATAGATCAGAACAAAAGTTATGTCGCAAAATTTGCAACCAATCTTTGTGAATTAC 358

Qy   1408 tttttttgcatctttaataattcttatatgttgtagagtagtgtttttaaataaattc 1467
          ||||| |||  ||| |  ||| | ||  ||||||||||||||||||||||||| || | |
Db    357 TTTTTGTTACTCATTTCCTGTGAGATCCCGGTGTGTTTTGTGATTTTTAAAAAATATACC 298

Qy   1468 aagtattttttaaaaaac 1486
          |||||||||||||||
Db    297 AAAATGCTTTTTCACAAAAC 279
```

```
RESULT   9
US-08-722-001-7
; Sequence 7, Application US/08722001
; Patent No. 5760054
; GENERAL INFORMATION:
;   APPLICANT: Thompson, Wayne J.
;   APPLICANT: Huff, Joel R.
;   APPLICANT: Nerenberg, Jennie B.
;   APPLICANT: Lee, Hee-Yoon
;   APPLICANT: Bell, Ian M.
;   TITLE OF INVENTION: ALPHA1C ADRENERGIC RECEPTOR ANTAGONISTS
;   NUMBER OF SEQUENCES: 35
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Merck & Co., Inc.
;     STREET: 126 Lincoln Avenue
;     CITY: Rahway
;     STATE: New Jersey
;     COUNTRY: United States of America
;     ZIP: 07065
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/722,001
;     FILING DATE:
;     CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/229,276
;     FILING DATE: 14-APR-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Appollina, Mary A.
;     REGISTRATION NUMBER: 34,087
;     REFERENCE/DOCKET NUMBER: 19169Y
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (908)594-3462
;     TELEFAX: (908)594-4720
;     TELEX: 138825
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1601 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: both
;     TOPOLOGY: both
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
US-08-722-001-7

  Query Match    2.4%;   Score 36.4;  DB 2;  Length 1601;
  Best Local Similarity   52.7%;   Pred. No. 0.39;
  Matches   79;  Conservative   0;  Mismatches   71;  Indels   0;  Gaps   0;
```

```
Qy    538 cgcgtccacggcctccagcccacgggacacggcgcgctggctgccgttatctgcaagcgctct 597
          || ||  | || || |  | ||| | ||  |||| | ||  | | |||| | | | | ||
Db   1107 CGAGGACGAGACCATCTGCCAGATCAACGAGAGCCGGCGTACGTGCTCTTCTCCGGCTCT 1166
```

Qy   598 ggccacccgtcctgtctggccctgctcatcctctgtgtcatctatattgtaagagacagtttat 657
          ||| ||| |||||| | || ||| ||| ||||| || | || | | || ||
Db   1167 GGGCTTCCTTCTACCTTGCCCTCTGGCCATCACCTGGTCTCATGTACTCCCGCGTCTACGTGGT 1226

Qy   658 ggaaaagaaacccagctggtctctcgcggtc 687
          | | |||||| | || | | || |||
Db   1227 GGCCAAGAGGGAGAGCGCGGGGGCCTCAAGTC 1256

RESULT 10
US-08-722-001-26
Sequence 26, Application US/08722001
Patent No. 5760054
GENERAL INFORMATION:
APPLICANT: Thompson, Wayne J.
APPLICANT: Huff, Joel R.
APPLICANT: Nerenberg, Jennie B.
APPLICANT: Lee, Hee-Yoon
APPLICANT: Bell, Ian M.
TITLE OF INVENTION: ALPHA1C ADRENERGIC RECEPTOR ANTAGONISTS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,001
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/229,276
FILING DATE: 14-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Appollina, Mary A.
REGISTRATION NUMBER: 34,087
REFERENCE/DOCKET NUMBER: 19169Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)594-3462
TELEFAX: (908)594-4720
TELEX: 138825
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1987 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-722-001-26

Query Match   2.4%;   Score 36.4; DB 2; Length 1987;
Best Local Similarity 52.8%;   Pred. No. 0.45;
Matches   79; Conservative   0; Mismatches 71; Indels   0; Gaps   0;

Qy   538 cgcgtccacggcctccacgcccacgggacacggggcgcgctggctgccgttatctgcagcgctct 597

Qy   598 ggccacccgtcctgtctggccctgctcatcctctgtgtcatctatattgtaagagacagtttat 657
          |||||| | ||| || | | ||||||||| |||| | || | | || ||
Db   1112 CGAGGACGAGACCATCTGCCAGATCAACGAGGAGCCGGGCTACGTGGTCTTCTCGGCTCT 1171

Qy   598 ggccacccgtcctgtctggccctgctcatcctctgtgtcatctatattgtaagagacagtttat 657
Db   1172 GGGCTTCCTTCTACCTTGCCCTCTGGCCATCATCCTGGTCTCATGTACTCCCGCGTCTACGTGGT 1231

Qy   658 ggaaaagaaacccagctggtctctcgcggtc 687

Db   1232 GGCCAAGAGGGAGAGCGCGGGGGCCTCAAGTC 1261

RESULT 11
US-08-722-001-27
Sequence 27, Application US/08722001
Patent No. 5760054
GENERAL INFORMATION:
APPLICANT: Thompson, Wayne J.
APPLICANT: Huff, Joel R.
APPLICANT: Nerenberg, Jennie B.
APPLICANT: Lee, Hee-Yoon
APPLICANT: Bell, Ian M.
TITLE OF INVENTION: ALPHA1C ADRENERGIC RECEPTOR ANTAGONISTS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,001
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/229,276
FILING DATE: 14-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Appollina, Mary A.
REGISTRATION NUMBER: 34,087
REFERENCE/DOCKET NUMBER: 19169Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)594-3462
TELEFAX: (908)594-4720
TELEX: 138825
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1997 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-722-001-27

Query Match   2.4%;   Score 36.4; DB 2; Length 1997;
Best Local Similarity 52.7%;   Pred. No. 0.45;
Matches   79; Conservative   0; Mismatches 71; Indels   0; Gaps   0;

Qy   538 cgcgtccacggcctccacgcccacgggacacggggcgcgctggctgccgttatctgcagcgctct 597
          |||||| | ||||| | ||||||||| |||| | || | | || ||
Db   1106 CGAGGACGAGACCACTCGCCAGATCAACGAGGAGCCGGGCTACGTGGTCATGTACTCGCGGCTCT 1165

Qy   598 ggccacccgtcctgtctggccctgctcatcctctgtgtcatctatattgtaagagacagtttat 657
          ||| ||| |||||| | || ||| ||| ||||| || | || | | || ||
Db   1166 GGGCTTCCTTCTACCTTGCCCTCTGGCCATCATCCTGGTCTCATGTACTCCCGCGTCTACGTGGT 1225

Qy   658 ggaaaagaaacccagctggtctctcgcggtc 687
          | | |||||| | || | | || |||
Db   1226 GGCCAAGAGGGAGAGCGCGGGGGCCTCAAGTC 1255

RESULT 12

US-08-722-001-11
Sequence 11, Application US/08722001
; Patent No. 5760054
; GENERAL INFORMATION:
; APPLICANT: Thompson, Wayne J.
; APPLICANT: Huff, Joel R.
; APPLICANT: Nerenberg, Jennie B.
; APPLICANT: Lee, Hee-Yoon
; APPLICANT: Bell, Ian M.
; TITLE OF INVENTION: ALPHAIC ADRENERGIC RECEPTOR ANTAGONISTS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,001
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/229,276
; FILING DATE: 14-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Appollina, Mary A.
; REGISTRATION NUMBER: 34,087
; REFERENCE/DOCKET NUMBER: 19169Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-3462
; TELEFAX: (908)594-4720
; TELEX: 138825
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2004 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-722-001-11

Query Match   2.4%;   Score 36.4; DB 2; Length 2004;
Best Local Similarity   52.7%;   Pred. No. 0.45; Mismatches 71; Indels 0; Gaps 0;
Matches   79; Conservative   0;

Qy   538 cgcgtccacggcctccagcccacgggacaacgggcgttattctgcagcgctct 597
          ||| || |||| || |||| |||||||||||||||
Db   1107 CGAGGACGAGACACATCGCCAGATCAACGAGGAGGACCGGCGGCTACGGTGCTCTTCTCGGCTCT 1166

Qy   598 ggcaccgtcctgtgtggccctgctcatcctctgtcatcatattgtaagagacaagtttat 657
          || || | || ||| ||||| || ||| || || |
Db   1167 GGGCTCCTTCTACCTGCCCATCAATCCTGGTCATGATACTGCCGGGTCTCACGTGGT 1226

Qy   658 ggaaaagaaaccaagctggtctctgccgtc 687
          || | |||| |||| |||||||
Db   1227 GGCCAAGAGGAGAGCCGCGGCGCCCTCAAGTC 1256

RESULT 13
US-08-424-424B-1
; Sequence 1, Application US/08424424B
; Patent No. 5759854
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.

; TITLE OF INVENTION: Neurotransmitter Transporter
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,424B
; FILING DATE: APRIL 21, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05363
; FILING DATE: MAY 25, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2485 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: cDNA
US-08-424-424B-1

Query Match   2.4%;   Score 36.2; DB 2; Length 2485;
Best Local Similarity   56.2%;   Pred. No. 0.59; Mismatches 53; Indels 0; Gaps 0;
Matches   68; Conservative   0;

Qy   517 caaggtcaacctcgtgaagatcgcggtccacggcctccagcccacgggacacgggcgctggc 576
          ||| | ||| |||| |||| |||| |||||||||||||||||||||||||| |||||
Db   2023 CAGCATCATCCAGCTGGGGTCAGCGCCCCCCGGCTACAGCGCGCCTGGATCAAGAGGAGGAGGC 2082

Qy   577 tgccgttatctgcagcgctctggcaccgtcctgtgtggccctgctcatcctctgtgtcat 636
          ||||| | ||| || ||| |||| | || || |||||||| ||||| ||| | ||
Db   2083 TGCCGAGCGCGCGTACCTGTATAATTCCCAACTGGCCATGGCACTCCTGATCACCCTCATCGT 2142

Qy   637 c 637
          |
Db   2143 C 2143

RESULT 14
PCT-US94-05363A-1
; Sequence 1, Application PC/TUS9405363A
; GENERAL INFORMATION:
; APPLICANT: LI, ET AL.
; TITLE OF INVENTION: Neurotransmitter Transporter
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE

; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05363A
; FILING DATE: SUBMITTED HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-118
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2486 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
PCT-US94-05363A-1

Query Match   2.4%;   Score 36.2; DB 6;   Length 2486;
Best Local Similarity 56.8%; Pred. No. 0.59;
Matches   68; Conservative   0; Mismatches 53; Indels   0; Gaps   0;

Qy 517 caaggtcaacctcgtgaagatcgvcgtccaacgvcctccacgvcaggagcaacggvcgctggc 576
Db 2024 CAGcATCATCCAGCTGGGGGGTCACGCCCCCCCGGCCTTACAGGCGCCTGGATCAAGGAGGAGGC 2083

Qy 577 tgccgttatctgcagcgcgctctggccacgvctcctgvcgccctgvctcatcctctgvtgvtcat 636
Db 2084 TGCCGAGGCGTTACCTGTATTTCCCCAACTGGCCCAcATGGCACTCCTGATCACCCTCAtCGT 2143

Qy 637 c 637
Db 2144 C 2144

RESULT 15
US-08-705-937-10/c
Sequence 10, Application US/08705937
Patent No. 5981841
GENERAL INFORMATION:
APPLICANT: Santino, Colleen G.
APPLICANT: Conner, Timothy W.
TITLE OF INVENTION: EARLY SEED 5' REGULATORY SEQUENCE
NUMBER OF SEQUENCES: 12
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,937
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Nicolas G. Barzoukas.
REGISTRATION NUMBER: 38,823
REFERENCE/DOCKET NUMBER: MOBT:018 (38-2(10694)A

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,937
FILING DATE:
CLASSIFICATION: 435
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carmen Rodriguez, Paralegal, Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210

; NAME: Janelle D. Waack,
; REGISTRATION NUMBER: 36,300
; REFERENCE/DOCKET NUMBER: MOBT:018 (38-2(10694)A
; NAME: Barbara S. Kitchell
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MOBT:018 (38-2(10694)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 787-1400
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1183 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-705-937-10

Query Match   2.4%;   Score 35.4; DB 4;   Length 1183;
Best Local Similarity 52.6%; Pred. No. 0.64;
Matches   72; Conservative   0; Mismatches 65; Indels   0; Gaps   0;

Qy 1352 acagtacagatcagaacctctgvtcccagvcataagvcatttggggaacctggatgvgvagtttttt 1411
Db 426 AAAGAATATATTAANNANATGTTAAAAATATATTTTTAACCCTCGTTCAATTTTTT 367

Qy 1412 ttttgvcatcttaataattcttatatgvttgvagagvtatgvtttaaaataaattcaagt 1471
Db 366 TTATGGATGCTAAANAGTGTAATAATTTAGTTTCAGTATGTCCTTAAATAATATTTCTTT 307

Qy 1472 atttttttaaaaacctt 1488
Db 306 TTGTATGGATAAACTATT 290

Search completed: April 23, 2000, 02:01:41
Job time: 1275 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2000, 01:41:21 ; Search time 42.52 seconds
(without alignments)
8761.430 Million cell updates/sec

Title: US-09-490-187-1
Perfect score: 1489
Sequence: 1 ggacctgcagcctcccaggt.........gtatttttttaaaaacttt 1489

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Match	Query Length	DB	ID	Description
	1	1489	100.0	1489	1	X23415	Human hAPO4-alpha
	2	1421.4	95.5	1496	1	V33362	Nucleotide sequenc
	3	1300.8	87.4	2185	1	X24978	Human TRAIN-R cDNA
	4	1268.4	85.2	1704	1	X3361	Nucleotide sequenc
	5	733	49.2	1678	1	X23413	Mouse mAPO4-alpha
	6	457.6	30.7	886	1	X23414	Mouse mAPO4-alpha
	7	457.6	30.7	942	1	X24977	Mouse TRAIN-R (lon
	8	360	24.2	371	1	V11422	Mouse secreted pro
	9	360	24.2	591	1	V11422	Human secreted pro
	10	325	21.8	599	1	X24976	Mouse mAPO4-gamma
	11	325	21.8	1133	1	X24979	Mouse TRAIN-R (sho
	12	120.2	8.1	791	1	X23416	Rat rAPO4-alpha DN
	13	90.4	6.1	181	1	V11423	Clone GJ156 encodi
	14	87.4	5.9	201	1	V86655	Human secreted Pro
	15	80.4	5.4	396	1	X23418	EST clone AX92. Ne
	16	39.6	2.7	4673	1	Q27189	Mouse mAPO4-beta D
	17	38.6	2.6	1558	1	T28625	P. yoelii SSP2 ant
	18	38.6	2.6	1560	1	T18828	Survival motor neu
	19	38.6	2.6	1582	1	T18831	Survival motor neu
	20	38.6	2.6	1582	1	T28829	Human survival neu
	21	38.4	2.6	326	1	O05546	Human survival mot
	22	38.2	2.6	4315	1	V80619	Human survival mot
	23	37.4	2.5	5805	1	X2028	Kidney injury asso
	24	36.8	2.5	1453	1	N91505	Borrelia burgdorfe
	25	36.6	2.5	761	1	X22851	Maize 33c BamHI fr
	26	36.6	2.5	761	1	V62663	LO-CD2a light chai
o	27	36.6	2.5	807	1	V62665	LO-CD2a antibody V
o	28	36.6	2.5	1602	1	Q7217	Humanised LO-CD2a
o	29	36.4	2.4	1978	1	Q7217	Human adrenergic r
o	30	36.4	2.4	1978	1	Q7212	Truncated human al
o	31	36.4	2.4	1997	1	T11624	Human truncated al
o	32	36.4	2.4	1998	1	T11600	Human alpha-1C adr
o	33	36.4	2.4	2004	1	Q7213	Human alpha-1C adr
o	34	36.4	2.4	2005	1	T11599	Human alpha-1C adr
c	35	36.2	2.4	2486	1	T09866	Human neurotransm
c	36	36	2.4	1193	1	V74563	Staphylococcus aur
c	37	36	2.4	2067	1	V99092	DNA methyltransfer
c	38	36	2.4	19440	1	V99129	DNA methyltransfer
c	39	35.6	2.4	761	1	Q71876	LO-CD2a VL coding
c	40	35.4	2.4	1183	1	T29006	Nucleotide sequenc
c	41	35	2.4	403	1	T29221	Parietaria allerge
c	42	35	2.4	845	1	O03998	Sequence complemen
c	43	34.8	2.3	1639	1	O63182	Alpha 1c adrenergi
c	44	34.8	2.3	1639	1	O62818	Genomic sequence e
c	45	34.8	2.3	1639	1	T03129	Alpha-1C adrenergi

RESULT 1

X23415

```
ID   X23415 standard; DNA; 1489 BP.
XX
AC   X23415;
XX
DT   18-JUN-1999  (first entry)
XX
DE   Human hAPO4-alpha DNA.
XX
KW   Tumour Necrosis Factor receptor; signal transducer molecule; TNF; APO4;
KW   developmental abnormality; gestational abnormality; prostate cancer;
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW   apoptosis; human; APO4-alpha; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             65..1336
FT                   /*tag= a
FT                   /product= "hAPO4-alpha"
XX
PN   WO9911791-A2.
XX
PD   11-MAR-1999.
XX
PF   04-SEP-1998; U:83393.
XX
PR   05-SEP-1997; US-924634.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Chaudhary PM;
XX
DR   WPI; 99:205191/17.
DR   P-PSDB; W93581.
XX
PT   New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT   useful for diagnosis and treatment of prostate cancer and
PT   developmental or gestational abnormalities
XX
PS   Example IV; Fig 7C; 1566p; English.
XX
CC   This invention describes isolated Tumor Necrosis Factor (TNF) family
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC   their active fragments. APO4 is useful for diagnosing prostate cancer
CC   by determining levels of APO4 in an individual. Prostate cancer can also
CC   be treated using APO4 selective binding agents linked to a therapeutic
CC   moiety. APO4 polypeptides are also useful for identifying selective
CC   binding agents. useful in diagnosis/treatment of disease by binding of
CC   agents to the polypeptide/active fragment which is extracellular, or
CC   expressed on the cell surface. The binding is preferably performed in
CC   vivo. APO4 polypeptides/ active fragments are also useful for screening
CC   for agonists and antagonists by binding and observing the change in APO4
CC   activity. Effective pharmacological agents useful in diagnosis or
CC   treatment of disease are also identified using APO4 polypeptides/active
CC   fragments and APO4 signal transducer molecules that specifically interact
CC   with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC   activity. The method is performed in vivo or in vitro. APO4 polypeptides
CC   are all useful as immunogens for preparing antibodies. APO8 is also
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma cell line
CC   MCF-7, and induced apoptosis.
XX
SQ   Sequence 1489 BP; 361 A; 366 C; 380 G; 382 T;
```

ALIGNMENTS

RESULT 1

```
Qy          1 ggacctgcagcctcccaggtggctgggaagaactctccaacaataaatacattgataag 60
```

Query Match   100.0%; Score 1489; DB 1; Length 1489;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1489; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Db      1 GGACTTGCAGCCTCCCAGGTGGCTGGGAGGAACTCTCCAAACAATAAATACAATTTGATAAG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 aaagatggtgcttaaaagtgctactagaacacaagaaaacgttttcactcctttagtatt 60

Db     61 AAAAGATGGGCTTTAAAAGTGCTACTAGAACAAGAGAAAAGGTTTTTCACTCCTTTAGTATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61 actaggctatattgtcatgtaaagtgactgtagacagcaagaatt 120

Db    121 ACTAGGCTATTTGTCATGTAAAGTGACTGTGAATCAGGAGACGTAGACAGCAAGAAATT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121 caggggatcggtctggaaactgtgttccctgcaacccagtgggaagttgtc 180

Db    181 CAGGGGATCGGTCTGGAAACTGTGTTCCCTGCAACCCAGTGGGAGTTGTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181 taaggaatgtggcttcggctatgggagtgacaagtgtgttgccggctgcacag 240

Db    241 TAAGGAATGTGGCTTCGGCTATGGGAGGATGCACAAGTGTGTTGCCGGCATGGGAGTTGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241 gttcaaggaggactgggcttccagaaatgcaagccctgtctggactggtgaa 300

Db    301 GTTCAAGGAGGAGACTGGGCTTCCAGAGAATGCACGGAGACTGTAGACACAAGAGAATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 ccgttcaagggaaagcaaaattgttcagccacccaccagtgatgcctgcc 360

Db    361 CCGCTTTCAGAAGGGCAAAATTGTTCAGCCACCACCAGTGATGCCTGCCTGGTGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361 aggatttttcaggagaacgacaaactgttcgcgcttcaagacaatgaagtgtgtgcttgtgg 420

Db    421 AGGATTTATAGGAGGAACGAAAACTTGTCGGCTTCAAGACATGAAGTGTGTGCCTTGTGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 agacctctcctcgctgttgcctgtctcatcctctgttctattgttaagagaatcgc 480

Db    481 AGACCTCTCCTCTCCTTACGAACCGCACTCGTGCCCAGTCAAGGTTGCAACCTTCGTGAAGATCGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481 gtccacggcggcctccacacggggacacgggcttatctgcgcgctctggc 540

Db    541 GTCCACGGCGGCCTCCACACCACGGGACACGGCCTGCTTCCACGGCCTGCTCCAGCCTCGGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541 cacggtcctgctgggcctgtctcactccattgtgttgttatggaagagaagttatgga 600

Db    601 CACGGTCCTCGTGCCCTGCCGTCAATCCTGTGTGTCATCTATTGTGTAAGAGAAGTTTATGGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601 aaggaaaccagctggtctctgcggtcacagacattcagtacacgagagactgtgtc 660

Db    661 AAAGAAAACCCAGCTGGTCTCTGCGGTCACAGACATTCAGTACACGAGAGACTGAGCTGTC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    661 gtgtttttgacagacctcacagtatatgcccacacaagacctgctgcacagtctgcccg 720

Db    721 GTGTTTTGACAGACCTCAGCTGCTTCAGGAATATGCCCACGTCCTGCGCCAGTGTGCCGCCG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    721 ctgcagcccaaccggcgactctggttcgttgtggggtgcattctggcagtctcaggc 780

Db    781 CTGCAGCCCCAACCGGCACTCTTGGTTCGTTGTGGGGTGCATTCTGCAGTCTTCAGGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    781 tgactcagttgcagaactgcgggtgcgcctgctctccccatccatgtgtctgttgaagaggc 840

Db    841 TGACTCAGTGCAGAACCTGCGCGTGCGCCTTCCTCCCATCCATGTGTGTGCTGTGAGGAGGC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    841 aagaaaaccgacccaaccgggactcctggttgttggggtgcattctgcaagccagtctcaagtc 900

Db    901 AAGAAAACGGACAGGCCCGGGAGAATGGTCCGGACTTTCTTCGGATCCCTCACCAGTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901 catctgtggccgagtttcagatgtcagatgtcagaatccacatgggtttgacaa 960

Db    961 CATCTGTGGCGAGTTTCAGATGTCGGACTTTCTTGGTGGCTTGAGGAGGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    961 cattctcttttgtgactcttatcctgaactccatcttcttcaatcc 1020

Db   1021 CATCTCTTTTGTGACTCTTATCCTGAACTCATTCTCTCAATCC 1080
          |||||||||||||||||||||||||||||||||||||||||||
Qy   1021

Db   1081 AGAACTTGAAAGCTCAAACGTCTTGGATTCAAATAGCAGTGACAGACTTTGGTGGTGGGGC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1081 agaacttgaaagtcaaacgtcttggatcaaatagcagtcaagactt 1140

Db   1141 TGTTCCAGTCCAGTCTCATTCCTGAAAACTTACAGCAAGTGACTTTATCTAGATATAA 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1141 tgttccagtcctccattcctgaaaacttacagcaagtctgattttatctagatat 1200

Db   1201 CAACACAACTGGTAGAATCAGCAATCACCATCAACTCACCATTAAGAGAAGCCACCTTAGA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1201 caacacactggtagaatcagcaccatcaactcaccaactatgagaagccagctaga 1260

Db   1261 TCAGGAGAGTGGCGCGTCTCATCCACCCAGCACTCAGACCCCCCCCAGGTAAGGCAGCG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1261 tcaggagagtggcgcgtctcatccacccagcaccaggtaaggcagcg 1320

Db   1321 ACTGGGTTCCCTGTGAACACAGCACTGACTTCAGTAGGACTTCTGTTCCCCAGCAT 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1321 actgggtttccctgtgaacacagcactgacttacagtagtctgttcccagcat 1380

Db   1381 AAGAATTTGGGGGAACCTGAGTAGTTTTTTTTTTTTGCATCTTTAATAATTTCTTATATGT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1381 aagatttggggaacctggatgagtttttttttttgcattctttaataatttctatatgt 1440

Db   1441 TGTAAGGTATGTTTTAAAATTACAATTGTATAGATTTTTTTAAAAACTTT 1489
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1441 tgtaagtatgttttaaaaattcaagtatttttttttaaaaaacttt 1489
```

RESULT 2

V33362

ID   V33362 standard; cDNA to mRNA; 1496 BP.
AC   V33362;
DT   02-DEC-1998 (first entry)
DE   Nucleotide sequence of human beta-OAF065.
KW   Human; beta-OAF065; stroma cell; antibody; inflammatory;
KW   cytokine-mediated disease; rheumatism; ulcerative colitis; ss.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
CDS   45..1316
               /*tag= a
               /product= "human beta-OAF065 protein"
sig_peptide   45..119
               /*tag= b
mat_peptide   120..1313
               /*tag= c
               /transl_except= (pos:711..713, aa= Pro)
               /transl_except= (pos:714..716, aa= Arg)
PN   WO9838304-A1.
PD   03-SEP-1998.
PF   26-FEB-1998; J00799.
PR   27-FEB-1997; JP-043143.
PA   (ONOY) ONO PHARM CO LTD.
PI   Fukushima D, Konishi M, Tada H;
DR   WPI; 98-481205/41.
DR   P-PSDB; W70387.
PT   Membrane polypeptide expressed by human stroma cells, and antibodies
PT   recognising it - for treatment of inflammatory and other
PT   cytokine-mediated diseases.
PS   Disclosure; Pages 40-41; 54pp; Japanese.
CC   This is the nucleotide sequence of the human beta-OAF065, used in
CC   the method of the invention. The process involves the use of peptides
CC   expressed by stroma cells, and its antibodies are used for in the
CC   prevention and treatment of inflammatory and other cytokine-mediated
CC   diseases such as rheumatism, ulcerative colitis.
SQ   Sequence 1496 BP; 388 A; 360 C; 372 G; 376 T;

Query Match   95.5%; Score 1421.4; DB 1; Length 1496;
Best Local Similarity   98.8%; Pred. No. 0;
Matches 1443; Conservative   0; Mismatches 16; Indels 2; Gaps 1;

```
Qy     27 gaagaacttcccaacaataaatacatttgataagaaagatgggcttaaaagtgctactag 86
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7 GTAGAACTTCTCCAACAATAAATACATTTGATAAGAAAGATGGGCTTAAAAGTGCTACTAG 66
```

```
Qy    87 aacaagagagaaacgttttccactcttttagttactaggctattgtcatgtaagtga   146
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67 AACACAGAGAAAACGTTTTTCACTCTTTTAGTTATTACTAGGCTATTGTCATGTAAAGTGA   126

Qy   147 cttgtgaatcaggagactgtagacacgacaagcaagaattcaggagatcggttctggaaactgtgttc   206
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 CTTGTGAAACAGGAGACTGTAGACACGACAAGCAAGAATTCAGGAGATCGGTTCTGGAAACTGTGTTC   186

Qy   207 cctgcaaccagtcgtgtgggccaaggccatggagttgtctcaaggaatgtggcttcggctatgggg   266
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 CCTGCAACCAGTCGTGTGGGCCAAGGCCATGGAGTTGTCTCAAGGAATGTGGCTTCGGCTATGGGG   246

Qy   267 aggatgcacagtcgtgtgggcggtgccacaggttcaaggcgggggcttccaga   326
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 AGGATGCACAGTCGTGTGACGTGCCGGTGCCACAGGTTCAAGGCGGGACTGGGCTTCCAGA   306

Qy   327 aatgcaagccctgctctggacgtgcgcaagtgttcaggagaggcaaattgttcag   386
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 AATGCAAGCCCTGTCTCGGACTCGCGCAGTGGTCAGGAGAGGCAAATTGTTCAG   366

Qy   387 ccaccagtgatgccatctgcgggactcctcctcctaacgaaccgc   446
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 CCACCAGTGATGCCATCTGCGGGGACTGTGCTTGCCCAGGATTTTATAGGAAGACGAAAACTTG   426

Qy   447 tcggcttttcaagacatggagtgtgtgccttgttggagaccctcctctcctaacgaaccgc   506
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 TCGGCTTTCAAGACATGGAGTGTGTGCCTTGTTGGAGACCCTCCTCTTCCTTACGAACCGC   486

Qy   507 actgtgccaccaaggtcaaacctcgtgaagatcgtcgtccacggccctccagcacgcgggaca   566
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 ACTGTGCCAGCAAGGTCAAACCTCGTGAAGATCGGTCCACGGCCCTCCAGCACGCGGACA   546

Qy   567 cggcgctggctggccgttatctgacagcgtcctggccaccgtcctgctgccctgctcatc   626
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 CGGCGCTGGCTGGCCGTTATCTGCAGCCGTCCTGGCCACCGTCCTGCTGCCCTGCTCATCC   606

Qy   627 tctgtgtcatctattgtaagagacagttattggaaaagaaaaccagctgttctctgtgcgt   686
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607 TCTGTGTCATCTATTGTGTAAGAGACAGTTTATGGAAAAGAAAACCAGTGTCTCTGTGGGT   666

Qy   687 cacagggacattcaagtacaacagaagactgaagctgtcgtgtttgacagaacctcaagctccacg   746
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667 CACAGGACATTCAGTACAACAGGCTCTGAGCTGTCGTGTTCTTGACAGAACCTCAAGCTCCAAG   726

Qy   747 aatatgccacagagctgctgccgtgactcagtgccagaactgcgggtccgg   806
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727 AATATGCCACAGAGCTGCTGCTGCGTGCCAGTGCCAGAACTGCGGGTCCGG   786

Qy   807 tgcgcttgctccatccatcgtcgtgttgaaggaggcctcgcagcccaaccaaccgggcgactcctg   866
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   787 TGCGCTTGCTCCCATCCATCCATGTCGTGCTGCTGGAGGAGGCCTCGCAGCCCCAACCCGGGGACTCTTG   846

Qy   867 gttgtgggtgcactctcgcagccagtctccaggcaagaacgcaggccagcccggggaga   926
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   847 GTTGTGGGGTGCGATTCTGCATTCTGCAGCCAGTCTCCAGGCAAGAACGGCCAGCCCGGGGAGA   906

Qy   927 tggtgcgcgacttcttttcttcggagatccctcacgcagtcttttttcaagatgcct   986
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   907 TGGTGCGCGACTTTCTTCGGATCCCTCACGCCAGTCTTTTTTCAGATGGCCT   966

Qy   987 ggccctctgatgcagaatccccatgggtcggtcgacaacatatctctttgtgacttctacctg   1046
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   967 GGCCTCTGATGCAGAATCCCCATGGGTCGGTCGACAACATCTCTTTTGTGACTCTTTATCCTG   1026

Qy  1047 aactcactgaggagacattcattctctccaatccaacgttggaagctcaaacgtcttttg   1106
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1027 AACTCACTGGAGGAAGACATTCATTCTCTCAAATCCAACTTGGAAAGCTCAAACGTCTTTGG   1086

Qy  1107 attcaaaatagacagtcaagaattggttcggtggggctgttccagtctcattctcaa   1166
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1087 ATTCAAATAGACAGTCAAGAATTTGGTTCGGTGGGGCTGTTCCAGTCTCATTCTCGAAA   1146
```

```
Qy  1167 actttacaagcagctactgatttatctagatatataacaacaactggtagaatcagcatcaa   1226
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1147 ACTTTACAGCAGCTACTGATTTATCTAGATATAAACAACAACACTGGTAGAATCAGCATCAA   1206

Qy  1227 ctcaggatgcactaactatgagaagccaagctgaatcacaggagagggagttgggcctgttcatcacc   1286
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1207 CTCAGGATG-CACTAACTATGAGAAGCCAGCTAGATCAGGAGAGGTGGGCTTATCATCACC   1266

Qy  1287 cagccacttcagacgttccctccagctaaggcacgacgactggtgttccctgtgaacaagcact   1346
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1267 CAGCCACTTCAGACGTCCCTCCAGTAAGGCAGGACGACTGGTGTTCCCTGTGAACAAGCACT   1326

Qy  1347 gacttacagtagaatcaagaacttgttccagcaataagatttggggaacctggatgagtt   1406
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1327 GACTTACAGTAGAATCAAGAAACTCTGTTCCCAGCAATAAGATTTGGGGAACCTG--ATGAGT   1384

Qy  1407 tttttttgtcatctttaatataattctttcttatatgtttgtagtagtatgttttaaaataaatt   1466
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1385 TTTTTTTTTGCATCTTTAATAATTTTCTTGTATGCTTGTGAGTAGTGTTTTAAAATAAATTT   1444

Qy  1467 caagtattttttttaaaaaact   1487
         ||||||||||||||||||||||
Db  1445 CAAGTATTTTTTTTTAAAAAACT   1465
```

RESULT　3

X24978
ID   X24978 standard; cDNA; 2185 BP.
AC   X24978;
DT   05-JUL-1999   (first entry)
DE   Human TRAIN-R cDNA.
KW   TRAIN-R; receptor; human; tumour necrosis factor receptor;
KW   agonist; antagonist; cancer; immunological disease; therapy;
OS   Homo sapiens..
OS   Homo sapiens..
FH   Key   Location/Qualifiers
FT   CDS   179..1432
FT   sig_peptide   179..253
FT   /*tag= a
FT   /*tag= b
FT   mat_peptide   254..1429
FT   /*tag= c
PN   WO9913078-A1.
PD   18-MAR-1999.
PF   11-SEP-1998; U19030.
PR   06-MAY-1998; US-084422.
PR   12-SEP-1997; US-058631.
PA   (BIOJ) BIOGEN INC.
PI   Hession C, Tschopp J;
PI   Hession C, Tschopp J;
DR   WPI; 99-229233/19.
DR   P-PSDB; W98146;
PT   New cysteine-rich tumor necrosis factor receptor
PS   Claim 1; Page 27; 30pp; English.
CC   The present sequence encodes a novel human cysteine-rich tumour
CC   necrosis factor receptor family member termed TRAIN-R (see W98146).
CC   It is a composite of 2 overlapping lambda gt10 clones (GJ159 and
CC   GJ158) from a Clontech human adult lung cDNA library.   Human
CC   TRAIN-R was also cloned from a second sequence subclone of a
CC   lambda gt10 cDNA (GJ156, see X24979).   Human TRAIN-R is expressed
CC   at low levels in every tissue and cell line tested thus far, with
CC   higher expression detected in heart, prostate, ovary, testis,
CC   peripheral blood lymphocytes, thyroid and adrenal gland.
CC   Cell death can be induced by administering an agent capable of
CC   inhibiting the binding of TRAIN-R to its ligand.   A claimed method
CC   of treating, or reducing, the advancement, severity or effects of
CC   an immunological disease in a mammal comprises administering a
CC   pharmaceutical composition which comprises a TRAIN-R blocking agent,
CC   e.g. soluble TRAIN-R.   TRAIN-R can be fused to an immunoglobulin to
CC   produce a fusion protein which may be targeted to various sites.
CC   It can be used in binding assays, and to identify antagonists and
CC   agonists.   Anti-TRAIN-R antibodies can be used to reduce the
CC   severity of an immune response or to treat cancer.   TRAIN-R
CC   blocking agents can also be used to reduce the severity or effects

CC   of an immunological disease (all claimed).
SQ   Sequence 2185 BP;   546 A;   551 C;   550 G;   538 T;

Query Match   87.4%;   Score 1300.8;   DB 1;   Length 2185;
Best Local Similarity   99.5%;   Pred. No. 0;
Matches 1305;   Conservative   0;   Mismatches   7;   Indels   0;   Gaps   0;

```
Qy       1 ggacctgcagcctcccgggtgggttgggggaagaaactctcaacaataataacacatttgataag 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115 GGACCTGCAGCCTCCCGGGTGGGTTGGGGGAAGAAACTCTCAACAATAATAACACATTTGATAAG 174

Qy      61 aaagatgtgcttaaaagtcgctactagaacaagaggagaaacgtttttcactcttttagtatt 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     175 AAAGATGTGCTTAAAAGTCGCTACTAGAACAAGAGGAGAAACGTTTTTCACTCTTTTAGTATT 234

Qy     121 actaggctcattgtcatgtaaagtgacttgtgaatcaggagaggagaatcagagacaagaatt 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     235 ACTAGGCTCATTTGTCATGTAAAGTGACTTGTGAATCAGGAGAGGAGACTGTGTAGACAGAGAATT 294

Qy     181 cagggacicgggtctggaaactgtgttccctgcaaccagtgtgggccagtgagcatgtagtgtc 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     295 CAGGGACTCGGTCTGGAAACTGTGTTCCCTGCAACCAGTGTGTGGGCCAGCATGTGGAGTTGTC 354

Qy     241 taaggaatgtgtggctctatgggggaggatgcacagtgcacagtgtgtgccggtgtgcacag 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     355 TAAGGAAATGTGGCTTGGCTTCGGCTATGGGGAGGATGCACAGTGTGTGACGTGCGCCGCTGCACAG 414

Qy     301 gttcaaggagaggactgggcttccacagaaatgcaagccctgtctggactgcgcagtggtgaa 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     415 GTTCAAGGAGGACTGGGCTTCCACAGAAATGCAAGCCCTGTCTGGACTGCGCAGTGGTGAA 474

Qy     361 ccgcttcaagaaggcaaattgttcagcaacgcactgtgatgccagtctgcgggggactgctggcc 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     475 CCGCTTCAAGAAGGCAAAATTGTTCAGCAACGCACTGTGGATGCCATCTGCGGGGGACTGCTTGCC 534

Qy     421 aggattttatggaagacgaaacttgtcggctttcaagacatggagtgtgtgtgccttgtgg 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     535 AGGATTTTATAGGAAGAGACGAAACTTGTCGGGCTTTCAAGACACGGAGTGTGTGGCCTTGTGG 594

Qy     481 agacccctcctcctctacgaaccgcactgtgccagcaaggtcaacctcgtgaagatcgc 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     595 AGACCCCTCTCCTCTTACGAACCGCACTGTGCCAGCAAGGTCAACCTTCGTTGAAGATCGC 654

Qy     541 gtccacacggcctcagccacgggacacgcgggcgccggttatctgcagcgctctggc 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     655 GTCCACACGGCCTCAGCCCACGGGACACGGGGCGCCGTTTATCTGCAGCGCTCTGGG 714

Qy     601 caccgtcctgctggccctgctcatctcctgtgtccatcctattgtaagagacagtttatga 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     715 CACCGTCCTGCTGGCCCTGCTCATCCTCTGTGTCATCTATTGTGTAAGAGACACAGTTTATGA 774

Qy     661 aaagaaaccagctgtctctcggtcaacggacattcagtacaacgagactgagctgtc 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     775 GAAGAAACCAGCTGTCTCTGGGTCTCCAACATTCAGTACAACGGCTTCAGTACTACAGCTTGTC 834

Qy     721 gtgtttgacaagacctcagctcacgaatatgcccacagagcctgtctgccagtgccgccg 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     835 GTGTTTTGACAGACCTCAGCTCCACGAATATGCCCACAGAGCCTGTCTGCCCAGTGCCCCCG 894

Qy     781 tgactcagtgcagacctcggggcgggcggttcgcgtcggcttcctcccatccatgctgctgtgaggaggc 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     895 TGACTCAGTGCAGACCTCGCGGGCCCTGGCCTTGCTGCCTCCCCATCCATGTGCTGTGAGGAGGC 954

Qy     841 ctgcagccccaaccgggggacttctggttgtggggtggctcttcagg 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     955 CTGCAGCCCCAACCGGGGGACTTCTGGTTGTGGGGTGGCTCTTCAGGC 1014

Qy     901 aagaaacagcagcccccacgcggggagatgtggttccgcacacttcttcggatcccctccacgccagtc 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1015 AAGAAACCCAGCAGCCCCACGCGGGGAGATGTGCGCGACACTTTCTTCGGATCCCCTCCACGCCAGTC 1074

Qy     961 catctgtgggcgagtttcagatgcctctggatgcagaatcccatgggggtggtgacaa 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1075 CATCTGTGGCGCGGAGTTTCCAGATGCCTGGCCTCTGGCCTGGGATGCAGAATCCCATGGGGTGGTGACAA 1134
```

Qy   1021 catctcttttttggactctttatcctggactcactggaggagaacattcattcctcaatcc 1080
Db   1135 CATCTCTTTTTTGGACTCTTTATCCTGGACTCACTGGAGGAGAACATTCATTCCTCAATCC 1194

Qy   1081 agaacttgaaagtcaaactcttttggattcaaatagcagtcagtcaaagaattcggttcggtggc 1140
Db   1195 AGAACTTGAAAGTCAAACTCTTTTGGATTCAAATAGCAGTCAGTCAAAGATTTTGGTTCGGTGGGGC 1254

Qy   1141 tgttccagtcccagtctcattctgaaaactttacgacgcagctactgattttatctcagatataa 1200
Db   1255 TGTTCCAGTCCCAGTCTCATTCTGAAAACTTTACGACGCAGCTACTGATTTATCTCAGATATAA 1314

Qy   1201 caacacactggttagaactcagcatcaactcaggaggatgcaactaactatgaggaagccagctaga 1260
Db   1315 CAACACACTGGTTAGAAATCAGCATCAACTCAGGAGGATGCCACTAACTATGAGGAAGCCAGCTAGA 1374

Qy   1261 tcaggagagtgggcgctgtcatccacccaagccactcaggacgtccctccaggta 1312
Db   1375 TCAGGAGAGTGGGCGGCGTCATCCACCCAAGCCACTCAGCACGTCCCTCCCACGGAA 1426

RESULT   4

ID   V33361   standard; cDNA to mRNA; 1704 BP.
AC   V33361;
DT   02-DEC-1998   (first entry)
DE   Nucleotide sequence of human alpha-OAF065.
KW   Human; alpha-OAF065; stroma cell; antibody; inflammatory;
KW   cytokine-mediated disease; rheumatism; ulcerative colitis; ss.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   CDS   45..1298
FT   /*tag= a
FT   sig_peptide   45..119
FT   /*tag= b
FT   mat_peptide   120..1295
FT   /*tag= c
FT   /product= "human alpha-OAF065 protein"
FT   /transl_except= (pos:711..713; aa= Pro)
FT   /transl_except= (pos:714..716; aa= Arg)
PN   WO9838304-A1.
PD   03-SEP-1998.
PF   26-FEB-1998; 1998.
PR   27-FEB-1997; JP-043143.
PA   (ONOY) ONO PHARM CO LTD.
PI   Fukushima D, Konishi M, Tada H;
DR   WPI; 98-481205/41.
DR   P-PSDB; W70386.
PT   Membrane polypeptide expressed by human stroma cells, and antibodies
PT   recognising it - for treatment of inflammatory and other
PT   cytokine-mediated diseases.
PS   Claim 5; Pages 31-32; 54pp; Japanese.
CC   This is the nucleotide sequence of the human alpha-OAF065, used in
CC   the method of the invention. The process involves the use of peptides
CC   expressed by stroma cells, and its antibodies are used for in the
CC   prevention and treatment of inflammatory and other cytokine-mediated
CC   diseases such as rheumatism, ulcerative colitis.
SQ   Sequence 1704 BP;   429 A;   426 C;   430 G;   419 T;

Query Match   85.2%;   Score 1268.4;   DB 1;   Length 1704;
Best Local Similarity   99.1%;   Pred. No. 0;
Matches 1275;   Conservative   0;   Mismatches   11;   Indels   0;   Gaps   0;

```
Qy      27 gaagaactctcaacaacaataataacacatttgataagaaagatgtgcttaaaagtgctactag 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7 GTAGAACTCTCAACAACAATAATAACACATTTGATAAGAAAGATGTGCTTTAAAAGTGCTACTAG 66

Qy      87 aacaagaggaaaacgttttcactctttttagtattactactaggctcatttgtcatgtaaagtga 146
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    67  AACAAGAGAAAAACGTTTTCACTCTTTTAGTATTACTAGGCTATTTGTCAGTGAAAGTGA   126

Qy   147  cttgtgttgaatcaggagaactgtagacagcaagaattcagggaaactgttgttc       206
          ||||||      |||||||||||||||||||||||||||||||||||||||||||
Db   127  CTTGTGAAACAGGAGACTGTAGACAGCAAGAATTCAGGGAAACTGTTGTTC           186

Qy   207  cctgcaaccagtgtgtgggccatgagtgagttgtctaaggaattgtgctatgggg       266
          ||||||||||||||||||||||   |||  |||||||||||||||||||||||||
Db   187  CCTGCAACCAGTGTGTGGGGCCAGGCATGGAGTTGTCTAAGGAATTGTGCTATGGGG      246

Qy   267  aggatgcaagtgtgtgggtgccggctgcacaggttcaaggagttcccaga           326
          |||||||||  ||||||| ||| |||||||||||||||||||   ||||||||
Db   247  AGGATGCACAGTGTGTGACGTGCCGGCTGCACAGGTTCAAGGGGCTTCCAGA          306

Qy   327  aatgcaagcccctgtctgactgcgcagttggtgaaccgctttcagggacaaatgttcaag  386
          |||||||||||||||||||  |||| | |||||||||||||||||  ||||   |||||||
Db   307  AATGCAAGCCCCTGTCTGGGACTGCCCGCAGTTGGTGAACCGGTTTCAGAAAGGCTTTCAG  366

Qy   387  ccaccagtgatgccatctgcgggggactgcctctgccaggatttttataggaagagcaaactg 446
          |||||||||||||| |||||||||||||||||| ||||||||||| |||||||||  |||||| 
Db   367  CCACCAGTGATGCCCATCTGCGGGGGACTGCCTTGCCAGGATTTTATAGGAAGACGAAACTTG  426

Qy   447  tcggcttcaagacaatggaagtgtgtgctgtggaagaaccgcttcctcctcctcacgaaccgc 506
          ||||||||||||| ||||||||||||| |||||| |||  | |||
Db   427  TCGGCTTCAAGACATGGAAGTGTGTGCCTGTGGAGAGGAGACTGGGGCTTCCAGA          486

Qy   507  actgtgtgccagcaagtcaacctcgtgactgcgcagttcggctctgaagatcctcacaggttc 566
          |||||| |||||||| ||||||||||||| | | |     ||| ||||||||||||| | |
Db   487  ACTGTGCCAGCAAGGTCAACCTCGTGACAGATCCCGGTCGAACTCTGAAGATCCTCACGGGACA 546

Qy   567  cggcgcgctggctgtattctgaagccggccccatcggtctctgctgggccctgctgtcatcc  626
          |||||| |||||||  | ||||  | | |   ||||| ||||| | |  | || |||||||||
Db   547  CGGCGCTGGCTGCGGTTATCTGCAGAGGCTTCTGCACGGCCTCCCACAGGGTGCATCC       606

Qy   627  tctgtgtcatcttattgttatggaaagaaacccagtctggtgttctgtcgcgt        686
          ||||||||||||||||||  | |   ||||||||||| ||| | |||| | |
Db   607  TCTGTGTCATCTATTGTAAGAACACATTTATTGAGAGAAACCCAGCTGGTCTCTGCGGGT    666

Qy   687  cacaggacattcagtacaacgagactgagctcgtcgttttgacagactcagtccacg     746
          |||||||||||| ||||||| | ||| | | ||||||||||||||| || || ||||
Db   667  CACAGGACATTCACTACAACGGCTCTGACTGTGTCGTTTTGACAGAGTCCAGCTCCACG     726

Qy   747  aatatgcccacagaagctgctgcacagtgcccagtgactcagtgcagagatgcaggggcgg   806
          ||||||||||||||||||||||||||||||||||||||||||||||  |    ||||| ||
Db   727  AATATGCCCACAGAAGCTGCTGCACAGTGCCCAGTGACTCAGTGCGCAGGGGCCCG        786

Qy   807  tgcgcttcgctcccatcgtgtctgtgtgaggaggccgtcagcccacccggccgggga      866
          ||||| ||||||||| | | ||||| | ||| ||||||| || | | |  | |||||
Db   787  TGCGCCTTCGCTCCCAACCATGTGTGTGAGGACCGTCGAGCCGTGCGACTCTTG         846

Qy   867  gttgtgtgggtgcattctgccagccagtcttcaggcaagaaacccagccgcgggggaga    926
          |||||||||||||||||||| ||||| |||| ||| |||||| | |||||| |||||||
Db   847  GTTGTGTGGGTGCATTCTGCAGTCTCGAGGCCAGAAAGCGCAAGAACGCGCCGGGGGAGA    906

Qy   927  tggtgccgacttcctcggatccccatctgtggggagttttcagatgcct            986
          ||||||||||| |||||||| || ||| ||||||| ||||||||||||
Db   907  TGGTGCCGACTTTCTTCGGATCCCATCCACGTGGGGGCGAGTTTTCAGATGCCT         966

Qy   987  ggccctcgatgcagaatcccatggttggtgacacacatctcatttcttttgtgactcttatcctg 1046
          || ||| |||||||||||||||||||||||||| |||||||| |||||||||||||||||||||||
Db   967  GGGCCTCTGATGCAGAATCCCATGGTTGGTGACAACATCTCTTTTCTTTTGTGACTCTTATCCTG   1026

Qy  1047  aactccactggagaagacattcattctctccaatccagaacttgaaagctcaacgtctttgg   1106
          |||| ||||||||||||||| |||||||||||||||||||  |||||||||||||||||||||
Db  1027  AACTGCACTGGAGAAGACAATTCATTCTCTCCAATCCAGAAGCTCAACGTCTTTGG          1086

Qy  1107  attccaaaatacgaagtcaaaaatttggttgttgtgggcgtgttccagtcctcattctgaa   1166
          ||||  |||||| |||||||| ||||||||| |||||||||||||||||| ||||||||||
Db  1087  ATTCAAAATAGACAGTCAAGATTTGGTTGGTGGTGGGCGTGTTCCAGTCTCATTCTGAAA     1146

Qy  1167  acttacgcagctactgattatctagatataacacactgtgtagaatcagcatcaa        1226
          ||||| |||||||||| ||| ||||||||||||||||||| ||||||||| ||||
Db  1147  ACTTTACACGCTACTGATTTACTAGATATAACAACACACTGTAGAATCACACATCAA       1206
```

```
Qy  1227  ctcaggatggcaactaactatagaagaccaagccagtctagatcaaggagagggggggctgtcatccacc 1286
          ||||||||||| |||||||| | ||||||||| |||||| ||||| ||||| |||||| | | ||||||||
Db  1207  CTCAGGATGGCGACTAACTAATCAGAAGACCAGCTAGATCCAGGAGGATCAGGAGTGGCGCGTATCATCCAC 1266

Qy  1287  cagccacctcagacgtccctccaagta 1312
          ||||||| |||||||| ||||||| ||
Db  1267  CAGCCACTCAGACGTGTCCTCCAGGAA 1292
```

```
RESULT   5
X23413

ID   X23413 standard; DNA; 1678 BP.
AC   X23413;
DT   18-JUN-1999 (first entry)
DE   Mouse mAPO4-alpha (long) DNA.
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW   developmental abnormality; gestational abnormality; prostate cancer;
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW   apoptosis; mouse; mAPO4-alpha; ss.
OS   Mus sp.
FH   Key              Location/Qualifiers
FT   CDS              72..1322
FT                    /*tag= a
FT                    /product= "mAPO4-alpha"
PN   WO9911791-A2.
PD   11-MAR-1999.
PF   04-SEP-1998; U18393.
PR   05-SEP-1997; US-924634.
PA   (UNIW ) UNIV WASHINGTON.
PI   Chaudhary PM;
DR   WPI; 99-205191/17.
DR   P-PSDB; W93579.
PT   New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT   useful for diagnosis and treatment of prostate cancer and
PT   developmental or gestational abnormalities
PS   Example IV; Fig 7A; 156pp; English.
CC   This invention describes isolated Tumor Necrosis Factor (TNF) family
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC   their active fragments. APO4 is useful for diagnosing prostate cancer
CC   by determining levels of APO4 in an individual. Prostate cancer can also
CC   be treated using APO4 selective binding agents linked to a therapeutic
CC   moiety. APO4 polypeptides are also useful for identifying selective
CC   binding agents, useful in diagnosis/treatment of disease by binding of
CC   agents to the polypeptide/active fragment which is extracellular, or
CC   expressed on the cell surface. The binding is preferably performed in
CC   vivo. APO4 polypeptides/ active fragments are also useful for screening
CC   for agonists and antagonists by binding and observing the change in APO4
CC   activity. Effective pharmacological agents useful in diagnosis or
CC   treatment of disease are also identified using APO4 polypeptides/active
CC   fragments and APO4 signal transducer molecules that specifically interact
CC   with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC   activity. The method is performed in vivo or in vitro. APO4 polypeptides
CC   are all useful as immunogens for preparing antibodies. APO4 is also
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma cell line
CC   MCF-7, and induced apoptosis.
SQ   Sequence   1678 BP;   371 A;   467 C;   466 G;   374 T;
```

```
Query Match         49.2%;   Score 733;   DB 1;   Length 1678;
Best Local Similarity   74.2%;
Matches  941;   Conservative   0;   Mismatches 325;   Indels   3;   Gaps   1;

Qy    42  aataaatacacttttgataaaaagtgcttactagaacaagagaaacgt               101
          ||||| ||||| ||||||||  | |||||||||||||| ||||| || |
Db    49  AATAAACACGTTTGGTGGAGACCCATGGCACTCAAGTCCTTACCTCTCACAGGACGGTGC   108

Qy   102  ttttcactctttagtattataataggctattgtcatgtaaagtgaatcaggag         161
          ||| |||||| || |  | || ||  |||| || | | || || | |  ||||||
Db   109  TCTTCGGCTG;CATTCTCTTCCTACTCCACCTGGGCACCTTGGAAGTGAGTTGCGAAACCGGAG  168
```

```
Qy    162 actgtagacagcaagaaattcaggatccggtcgtctggaaactgtgttccctgcaaccagtgtg 221
              |||| ||| |||||||||||||| |||| || |||||||||||||||||||||||||||||| ||
Db    169 AATTGCAGGCAGCAGGAATTCAAGGATCGATCTGGAAACTGTGTTCCCTGCAACCAGTGTGCG 228

Qy    222 ggccaggcaatsggaagtgtgtctaaaggaagtgtatgtgggaggaagtgtg 281
              ||||| || |||||||| |||||| ||||| | ||||||||||||||||||||
Db    229 GACCTGGCCATGGAGGAGTTGCCAAGGAATGTGGCCTGCAAACAGTGTG 288

Qy    282 tggcctgcccggctgcacagttcaaggagagagactgggggctcccagaaatgcaagccctgtc 341
              ||| ||| || ||||||| ||||||||||||||| |||||||| ||||| | ||||||| |||
Db    289 TGGCCTGCAGCCGCACCGGTTCAAGGAGAAGACTGGGGGTTTCCAGAAGTGTAAGCCATGTG 348

Qy    342 tggactgcgcagtggtgaaccgctcttcagagaggcaccacagtgatgcca 401
              ||||||| || |||||| || || |||| |||||| ||||| | ||| ||||
Db    349 CGGACGTGCGCGGTGGTGAACCGCTCCTTCAGAGAGGGCCAACTGCTCACACACCAGGCTG 408

Qy    402 tctgtcggggggactgctgttgccaggggacttttatgggaggagacgaacttgttcggcttcaagaca 461
              ||||||| ||||| |||| || ||||||||| |||| | ||||||||||||||| || ||||||||| ||
Db    409 TCTGCGGGGGGACTGCCTGCCCAGGATTTTACCGGGAAGACCAAAACTGGTTGGTTTCAGACA 468

Qy    462 tggagtggtgtgccttgtggagaccccactgtgtgccagcaagg 521
              |||||| ||| |||| ||||||| |||||||||||||||||||||
Db    469 TGGAGTGGTGTGCCCTGTGGGAGACCCGTTCAAGGAGAAGACTGGGGTTTTCACATGTG 528

Qy    522 tcaacctgtgaagtcgttccacggctcttcagagaggcaccacagtggctgccg 581
              |||||| |||||| || ||||||||| ||| ||| |||||| ||||||||||||||
Db    529 TGAACCTTGTGAAGATCTCCCACACGGGACAGGCTCCTTCAGACTCAACGGGCCG 588

Qy    582 ttatctgcaacgtcttctgctgtgctcctgtgctcatcttgtgtcatctatt 641
              |||||||||| ||| |||||||| ||||| ||||||| |||||||||| |||||
Db    589 TCATCTGCAGTGCTCTGGCCACGGTCGTTGTGCTCACTTGTGTCTCATCTACT 648

Qy    642 gtaagagacacagtttatggaaagaaagaaacccagctggttctctgtcggttcacaggacattcagt 701
              | ||||| |||||| | || |||||||| || || |||||| || ||||||| ||| || |||||||||
Db    649 GCAAGAGGCAGTTCAATGGAGAAGAAGAAAACCCAGCTGGTTCTCTGGGATCCACAGGACATTCAGT 708

Qy    702 acaacgagactgagctcgtcgttgtttgacagaacctcagctcccacacagag 761
              |||||| |||||||| ||||| || |||||||||||| ||| || |||||||||
Db    709 ACAATGGCTGGACTGTCGAGCTGTCATGCTTTGACCACGTGCCCATAGAG 768

Qy    762 cctgctgccagtgccgtcgactcagtgcagctgcggggcccgtgccttgtctccccat 821
              |||||| || ||||| |||||||||||| || ||| || |||||| ||||||||||||
Db    769 CATGCTGTCAGTATCACCGGGGACTCAGCCCCCAAATGTATGGGCCTGTTCACCTGATTCCGT 828

Qy    822 ccatgtgctctgaggagggcctgcagcccaacccggcgactcttggttgttgggtgcatt 881
              ||||||| ||||| || |||||| |||||| || ||||| ||||| || |||| || ||||
Db    829 CCTTGTGCTCTGAAGAGGCCCGCCAGCTCCCGAGCTGCGTCTTGGGCGTGTGGGCGTTT 888

Qy    882 ctgcagccagtcttcaggcaagaaacgccccagccggcggagatggttggccgagactttct 941
              ||| || ||||| ||||||||| || || |||||| |||||||| ||||| ||||| ||||||
Db    889 CTCCCACTACCCTCCAGGAGAGAAACCGGCTCTGTGGGGAACACGAATGGCCAGCCTTCT 948

Qy    942 tcggatcccccaagcagtcaatcgtctggccagtttcacaatgtctggcctctgatgcaga 1001
              |||||| || |||||||| ||||||||||||| || |||||| |||||||||||||||||||
Db    949 TTGGGTCTGTGTTTCCGTTCCACTCGGCGGAGTTTTCTGATGGCCTGCCTCTGATCGCAGA 1008

Qy   1002 atccaatgggtggtgacacatctcctttttgtgactctttatcctgaactcactggagaag 1061
              ||||| ||||| ||||||||| |||||||||||| ||| |||| ||||| |||||||||||
Db   1009 ATCCCTCTGGGCGGGTG---ACAGCTCTTCTGTGACTTTTAACCCTGAACTCACTGGGAAG 1065

Qy   1062 acaattcattctcaatccaagctcaacgtctctttggattcaaatagcagtc 1121
              |||||| |||| ||||| || ||||| || ||||||||| || |||| | |||
Db   1066 ATACCAATTCCCAATCCAGAAAACACAGCAGCCAGAGTCTTCGAATTCCAGTGCGGGCC 1125

Qy   1122 aagatttgttggggggtgttcaggtcattctgaaaacttttacacagcta 1181
              ||||| |||||| ||||| ||||| ||| ||||||| |||||||||||| ||
Db   1126 AGGAATCTGGCTGGTGGGGACAGCTGCTCTTGGGAAGTTTTCAGAATTCTACTGACT 1185

Qy   1182 ctgatttattcctagatataacacactggtacagcatctggtacagagacgctagctcagga 1241
              || ||| |||||| | |||||||||||||||| ||| || ||||||||| || |||||| |||
Db   1186 CACCTAGACACATGGTCGACACTGGTACAGTCTTCAGGAGTTGGGATGCTCAA 1245

Qy   1242 ctatgagaggccagctagatcaggagagtgcgcgtctcatccaccagccagccaactcagacgt 1301
              |||| |||||||||||||||||| ||| ||||||| || ||||||||||||||||||||||||||
Db   1246 GGACTCCAAGCCAAGGAGGCTTGGGAAGACAGCAGGGAAAACCTGAATTCTAGCCATGCCACAG 1305

Qy   1302 ccctccaagg 1310
              ||| ||||||
Db   1306 CCTTCCAGG 1314
```

```
RESULT  6
X23414
ID   X23414 standard; DNA; 886 BP.
AC   X23414;
DT   18-JUN-1999 (first entry)
DE   Mouse mAPO4-alpha (short) DNA.
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW   development;abnormality; gestational abnormalitiy; prostate cancer;
KW   APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; disease;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW   apoptosis; mouse; mAPO4-alpha; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   CDS             72..716
FT                   /*tag= a
FT                   /product= "mAPO4-alpha short"
PN   WO9911791-A2.
PD   11-MAR-1999.
PF   04-SEP-1998; U18393.
PR   05-SEP-1997; US-924634.
PA   (UNIW ) UNIV WASHINGTON.
PI   Chaudhary PM;
DR   WPI; 99-205191/17.
DR   P-PSDB; W93580.
PT   New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT   useful for diagnosis and treatment of prostate cancer and
PT   developmental or gestational abnormalities
PS   Example IV; Fig 7B; 156pp; English.
CC   This invention describes isolated Tumor Necrosis Factor (TNF) family
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC   their active fragments. APO4 is useful for diagnosing prostate cancer
CC   by determining levels of APO4 in an individual. Prostate cancer can also
CC   be treated using APO4 selective binding agents linked to a therapeutic
CC   moiety. APO4 polypeptides are also useful for identifying selective
CC   binding agents, useful in diagnosis/treatment of disease by binding of
CC   agents to the polypeptide/active fragment which is extracellular, or
CC   expressed on the cell surface. The binding is preferably performed in
CC   vivo. APO4 polypeptides/ active fragments are also useful for screening
CC   for agonists and antagonists by binding and observing the changer in APO4
CC   activity. Effective pharmacological agents useful in diagnosis or
CC   treatment of disease are also identified using APO4 polypeptides/active
CC   fragments and APO4 signal transducer molecules that specifically interact
CC   with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC   activity.The method is performed in vivo or in vitro. APO polypeptides
CC   are all useful as immunogens for preparing antibodies. APO4 is also
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma cell line
CC   MCF-7, and induced apoptosis.
SQ   Sequence    886 BP;    204 A;    245 C;    247 G;    190 T;

Query Match           30.7%;  Score 457.6;  DB 1;  Length 886;
Best Local Similarity  81.6%;
Matches  529;  Conservative  0;  Mismatches 119;  Indels  0;  Gaps  0;

Qy     42 aataaataacatttgataagaaagatggcttaaagtgctactagaacaagaagaaacgt 101
              ||| || | || | || || ||| ||||| |||||| || ||||||| ||||||||| ||
Db     49 AATAAAACACGTTTGGTGAGAGACCCATGGCACTCAAGGTCCTACCCTACACAGGACGGTGC 108

Qy    102 ttttcactctttagtatattactaggctatttgtcatgtaaagttgaatcaggag 161
              || | ||||| | || |||| |||| |||| || || ||||||| |||||| ||||
Db    109 TCTTCGCTGCCCATTCTTCCTACTCCACCTGGCATGTAAAGTGAAAGTGGAAACCGGAG 168
```

RESULT 7

```
ID   X24977 standard; cDNA; 942 BP.
AC   X24977;
DT   05-JUL-1999   (first entry)
DE   Mouse TRAIN-R (long form) cDNA.
KW   TRAIN-R; receptor; mouse; tumour necrosis factor receptor;
KW   agonist; antagonist; cancer; immunological disease; therapy;
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   CDS             101..745
FT                   /*tag= a
CC   New cysteine-rich tumor necrosis factor receptor
PN   WO9913078-A1.
PD   18-MAR-1999.
PF   11-SEP-1998; U19030.
PR   06-MAY-1998; US-084422.
PR   12-SEP-1997; US-058631.
PA   (BIOJ) BIOGEN INC.
PI   Hession C, Tschopp J;
DR   WPI; 99-229238/19.
DR   P-PSDB; W98145.
PS   Claim 1; Page 26-27; 30pp; English.
CC   The present sequence encodes a novel murine cysteine-rich tumour
CC   necrosis factor receptor family member termed TRAIN-R (long form)
CC   (see W98145). Murine TRAIN-R is expressed at high levels in brain
CC   and lung, and at lower levels in liver, skeletal muscle and kidney.
CC   Cell death can be induced by administering an agent capable of
CC   inhibiting the binding of TRAIN-R to its ligand. A claimed method
CC   of treating, or reducing, the advancement, severity or effects of
CC   an immunological disease in a mammal comprises administering a
CC   pharmaceutical composition which comprises a TRAIN-R blocking agent,
CC   e.g. soluble TRAIN-R (see also W98144). TRAIN-R can be fused to an
CC   immunoglobulin molecule to produce a fusion protein which may be
CC   targeted to various sites. It can be used in binding assays, and
CC   to identify antagonists and agonists. Anti-TRAIN receptor
CC   antibodies can be used to reduce the severity of an immune response
CC   or to treat cancer. TRAIN-R blocking agents can be used to reduce
CC   the severity or effects of an immunological disease (all claimed).
SQ   Sequence 942 BP;  219 A;  264 C;  258 G;  200 T;
```

```
Query Match             30.7%;  Score 457.6;  DB 1;  Length 942;
Best Local Similarity   81.6%;  Pred. No. 4.7e-121;
Matches 529;  Conservative  0;  Mismatches 119;  Indels  0;  Gaps  0;
```

```
Qy     42 aataaatacatttgtgataaggaagatggcttaaagtgctactagaacaagagaaacgt 101
          ||||| || |  ||||| || |||| ||| |||| |||||||||||  |||||||||||
Db     78 AATAAACACGTTTGGGTAGAAGGCCAAGGCACTCAAGGTCCTTACCCTCTACACAGGGTGC 137

Qy    102 ttttcaactcttttagtattactaggctcattcaatgtaaagtggactgtgaatcaaggg 161
          |||| ||||||||| ||||| || ||||||||| |||||| ||| || | ||| |||||||
Db    138 TCTTCGCTGCCACATTCTCTTCCTACTCCACCGGCATGTAAAGTGAGTTGCGGAAACCGGAG 197

Qy    162 actgtagacagcaagagaatcagggatcggtctgaaactgtgttccctgcaaccagtgtg 221
          |||| |||| |||||| ||| |||| ||| ||||||| |||||||| || ||||||||||||
Db    198 ATTGCCAGGCAGCAGGAATTCAAGGAATCGAACTGTGTGTTCCTCTGCAAAACAGTGCG 257

Qy    222 ggccaggcagtggagttgtctaaggaatgttggtcttcaagacctgtc 281
          |||||||||| |||| ||||| ||||||||||| | ||||||| ||| |
Db    258 GGCCAGGCAGGTGGAGTTGTCCAAGGAAATGGGCTTCGGCTATGGGGAGGATGCACAGTGTG 317

Qy    282 tggctgcctgaagatcgctccaggaggactggggcttcaagaaatgcaagccctgtc 341
          |||| ||||| ||||||||||||||| ||||| || ||| |||| |||||| ||| |||
Db    318 TGGCCTGCAAGGCCCACCGGTTCAAGGAAGACTGGGGTTTCCAGAAGTGTAAGCCATGC 377

Qy    342 tcaacctcgtgaagatcgctccaacggagccctctcaaccactgtacaccagtgg 401
          ||||| | ||||| |||||| || || |||  |||| |||  | || || |||||||
Db    378 CGGACTGTCGCGCTGGTGAACCGGTTTCAGAGGCCAACCAGTGATGCTGCCA 437

Qy    402 tctgcgggggactgcttgccaggagattttataggaagagagaaactgctcacacaccagtgatgctg 461
          |||| ||||| ||| || || |||  |||| || || ||||| |||||| |||||||||||  | ||
Db    438 TCTTCGCTGCCACATTCTCTTCCTACTCCACCGGCATGTAAAGTGAGTTGCGGAAACCGGAG 497

Qy    462 tggagtgtgtgccttgtgggagaccctcctctcctcctacgaaccgcactgtgccagcaagg 521
          |||||||| || |||||||| ||||||| ||| |||| ||  ||| || ||| || |||||||
Db    498 TGGAGTGTGCGCCCTGCGGGAGACCCCACCTCCCTACGAAACCACACTGTACCAGCAAGG 557

Qy    522 tcaacctcgtgaagatcgctccaacggaccctcctcaccagtgg 581
          ||||||| ||||| |||| ||||  ||| ||||| ||||||||||
Db    558 TGAACCTTGTGAAGATCTCCTCCACCGGTCTTCCAGCCTCTGCGG 617

Qy    582 ttatctgcagcgctctggccaccgtcctggcctgcctccatcctctgttgtcatctatt 641
          |||| || |||||||||| || ||| ||| ||| ||||| ||| |||| ||| |||||||
Db    618 TCATCTGCAGTGCCTCTGGCCGCTCTGGCCTGCCTCATCCTGTGTGTGTCATCTACT 677

Qy    642 gtaagagacagtttatggaaaagaaacccagctgttctctgcggtcac 689
          ||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    678 GCAAGAGGCAGTTCATGGAGAAGAAACCCAGCTGTAAGCTCCATCCC 725
```

RESULT 8

```
ID   V11422 standard; cDNA; 371 BP.
AC   V11422;
DT   23-JUL-1998   (first entry)
DE   Human secreted protein clone AX92_3 cDNA 5'-end.
KW   Secreted protein; prevention; treatment; gene therapy; ds.
OS   Homo sapiens.
PN   WO9801554-A2.
PD   15-JAN-1998.
PF   07-JUL-1997; U11876.
PR   09-JUL-1996; US-677231.
PA   (GEMY ) GENETICS INST INC.
PI   Bowman M, Evans C, Jacobs K, LaValle ER, McCoy JM,
PI   Merberg D, Racie LA, Spaulding V, Treacy M;
DR   WPI; 98-110230/10.
DR   P-PSDB; W58844.
```

CC   Secreted proteins and polynucleotides encoding them - useful to
PT   prevent, treat and ameliorate medical conditions
PS   Claim 15; Page 57; 93pp; English.
PN   V11422-V11424 encode fragments of a novel secreted protein derived from
CC   clone AX92_3 which was isolated from a human adult testes cDNA library.
CC   The protein can be used to prevent, treat or ameliorate a medical
CC   condition, while the polynucleotides can be used for gene therapy.
SQ   Sequence 371 BP;   83 A;   96 C;   107 G;   82 T;

Query Match   24.2%; Score 360; DB 1; Length 371;
Best Local Similarity   97.8%; Pred. No. 2.1e-93;
Matches   363; Conservative   0; Mismatches   8; Indels   0; Gaps   0;

Qy   192 ctggaaactgtgttccctgcaacaagtgtgggcaggcaggcgaagttgtctaaggaatgtg 251
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1 CTGGAAACTGTGTTCCCTGCAACAAGTGTGGGCCAGGCAGGCGAGTTGTCTAAGGAATGTG 60

Qy   252 gcttccggctatgggaggatgcacagtgtgtggcgtccggctgacacaggttcaaggagg 311
          ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   61 GCTTCCGGCTATGGGAGGATGCACAGTGTGTGGCGTCCGGCTGACACAGGTTCAAGGAGG 120

Qy   312 actgggggcttccagaaatgcaagccctgtctggactgagcagtggaaccgctttcaga 371
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   121 ACTGGGGGCTTCCAAAAATGCAAGCCCTGTCTGGACTGAGCCAGTGGTGAACCGCTTTCAGA 180

Qy   372 aggcaaattgttcagccaacagtgatgccatctgcgggggactgcttgccaggatttata 431
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   181 AGGCAAATTGTTCAGCCACCAGTGATGCCATCTGCGGGGGACTGCTTGCCANGATTTATA 240

Qy   432 ggaagacgaaacttgtccggcttcaagacatgagtggtgtgtgccaggttgggaaccctctc 491
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   241 GGAAGACNAAACTTGTGTCGGCTTCAAAACATGAGAGTGTGTGCCTTGTGGGAAACCCTCTC 300

Qy   492 ctcctacgaaccgcacctgtgccaggagtcaacctcgtgaagaatcggtccacgggcct 551
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   301 CTCCTTACGAACCGCACCTGTGCCAGCAAGGTCAACCTCGTGAAAATNCGTCCACGGCCT 360

Qy   552 ccagccacgg 562
          ||||||||||
Db   -361 CCACCCCACGG 371

Query Match   21.8%; Score 325; DB 1; Length 591;
Best Local Similarity   78.0%; Pred. No. 2.8e-83;
Matches   391; Conservative   0; Mismatches 110; Indels   0; Gaps   0;

CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC   their active fragments. APO4 is useful for diagnosing prostate cancer
CC   by determining levels of APO4 in an individual. Prostate cancer can also
CC   be treated using APO4 selective binding agents linked to a therapeutic
CC   moiety. APO4 polypeptides are also useful for identifying selective
CC   binding agents, useful in diagnosis/treatment of disease by binding of
CC   agents to the polypeptide/active fragment which is extracellular, or
CC   expressed on the cell surface. The binding is preferably performed in
CC   vivo. APO4 polypeptides/ active fragments are also useful for screening
CC   for agonists and antagonists by binding and observing the changer in APO4
CC   activity. Effective pharmacological agents useful in diagnosis or
CC   treatment of disease are also identified using APO4 polypeptides/active
CC   fragments and APO4 signal transducer molecules that specifically interact
CC   with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC   activity. The method is performed in vivo or in vitro. APO polypeptides
CC   are all useful as immunogens for preparing antibodies. APO4 is also
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma cell line
CC   MCF-7, and induced apoptosis.
SQ   Sequence 591 BP;   148 A;   145 C;   175 G;   123 T;

Qy   38 caacaataaatacaatttgatataagaaagatggctcttaaaagtgcttactagaacaagagaa 97
          || || |||||| ||||||| || |||||||||||||||||||||| ||||||| |||
Db   68 CAGGAATAAAACACGTTTGGTGAGAGCCAAGGCCAGGGCACTCAAGGTCCTCTACCTGTACAGGACG 127

Qy   98 acgttttcactctttttagtattaatactaggctatttgtcatgtaaagtaacttgtgaatca 157
          ||| | ||||| || ||| | |||| | || || |||| | || ||||||||| ||
Db   128 GTGCTCTTCGGCTGCCATTCTTCTCCTACTCCACCTGGCATGTAAAGTGAAGTTCGGAAACC 187

Qy   158 ggagacttgtagacagcaagaaattcaaggatcggtctgtggaaactgtgttccctgaaccag 217
          ||| |||||| ||| | |||||||||||||| | ||| ||||| ||||||||| ||
Db   188 GGAGAGTTGCAGGCAGCAGAAAATTCAAGGANCGGATCTGTCTCGTTGCAAACAG 247

Qy   218 tgtgggccaggccatggagtgttgtctaaggaaatgtggcttcggctatgggaggatgcacag 277
          |||||| |||| ||| ||| | ||||| |||||| |||||||| |||| |||| |||||||
Db   248 TGCCGGACCTGGCATGGAGTGTTGTCAAGGAAATGTGGCTTCGGCTATGGGGGGATGCACAG 307

Qy   278 tgtgtggcgtgcggctgcacaggttgtgcacaggttcaagaaaaatgcaagcc 337
          ||||||| | |||||||| || ||||||||| |||||| |||||||||||
Db   308 TGTGTGGCCTGCGGCCGCACCGGTTCAAGGAAGAAGACTGGGGTTTCCAGAAAGTGTAAGCCA 367

Qy   338 tgtctggactgcggagtggtgaaccgcttcagcacccaagtgat 397
          |||| || | |||||| ||| ||||||||||| | |||||| |
Db   368 TGTCGGGACTGTGGCGCTGGTGGAACCGCTTTCAGAGGGCCAACTGGACCACACAGTGTAT 427

Qy   398 gccatctcgggggactgcttgccaggatttatatattagggaagacgaaacttgtcggcttcaa 457
          ||| |||||||||||||||| ||| || ||||| |||| |||| ||||||| ||||||| ||
Db   428 GCTGTGCTGCCGGGACTGCCTGCCAGGATTTTACCGGAGACCAAAACTGCTTGGTGTTTTCAA 487

Qy   458 gacatggaagtgtgtggagacctccttcctcctaccgaaccgcaactgtgtccagc 517
          |||| ||| |||||| ||||||||| | ||||| |||||| || || |||||||
Db   488 GACATGGAGTGTGTGGGCTGGAGACCCCCTTACCGGAGACCAACACTGTTGGTGCAGC 547

Qy   518 aaggtcaacctgtgaaagatc 538
          ||||| |||||||||||||| |
Db   548 TGTGCCAAGTGGGACAGACC 568

CC   receptor polypeptides; APO4, APO6, APO8 and APO9 or their active
PT   This invention describes isolated Tumor Necrosis Factor (TNF) family
PS   Disclosure; Fig 7E; 156pp; English.
PT   developmental or gestational abnormalities
DR   New Tumor Necrosis Factor family receptor polypeptides and ligands -
PI   useful for diagnosis and treatment of prostate cancer and
PA   (UNIW) UNIV WASHINGTON.
PR   05-SEP-1997; US-924634.
PD   04-SEP-1998; U18393.
DT   11-MAR-1999.
PN   WO9911791-A2.

FT   CDS   95..547
FT   /tag= a
FT   /product= "mAPO4-gamma"
DR   WPI; 99-205191/17.
P-PSDB; W93583.
OS   Mus sp.

```
FH   Key              Location/Qualifiers
FT   CDS              104..556
FT                    /*tag= a
PN   WO9913078-A1.
PD   18-MAR-1999.
PF   11-SEP-1998; U19030.
PF   06-MAY-1998; U08422.
PR   12-SEP-1997; US-058631.
PA   (BIOJ ) BIOGEN INC.
PI   Hession C; Tschopp J;
DR   WPI; 99-229238/19.
DR   P-PSDB; W98144.
PT   New cysteine-rich tumour necrosis factor receptor
PS   Claim 1; Page 26; 30pp; English.
CC   The present sequence encodes a novel murine cysteine-rich tumour
CC   necrosis factor receptor family member termed TRAIN-R (short form)
CC   (see W98144). This putative natural soluble form of murine TRAIN-R
CC   may inhibit signalling by the full-length TRAIN-R (see W98145).
CC   Murine TRAIN-R is expressed at high levels in brain and lung, and
CC   at lower levels in liver, skeletal muscle and kidney. Cell death
CC   can be induced by administering an agent capable of inhibiting the
CC   binding of TRAIN-R to its ligand. A claimed method of treating, or
CC   reducing, the advancement, severity or effects of an immunological
CC   disease in a mammal comprises administering a pharmaceutical
CC   composition which comprises a TRAIN-R blocking agent, e.g. soluble
CC   TRAIN-R. TRAIN-R can be fused to an immunoglobulin molecule to
CC   produce a fusion protein which may be targeted to various sites.
CC   It can be used in binding assays, and to identify antagonists and
CC   agonists. Anti-TRAIN receptor antibodies can be used to reduce the
CC   severity of an immune response or to treat cancer. TRAIN-R
CC   blocking agents can be used to reduce the severity or effects of an
CC   immunological disease (all claimed).
SQ   Sequence 599 BP;       149 A;      147 C;      180 G;     123 T;

Qy   Query Match     21.8%;  Score 325;  DB 1;  Length 599;
     Best Local Similarity   78.0%;
     Matches 391;  Conservative   0;  Mismatches 110;  Indels   0;  Gaps   0;

Qy      38 caacaataataacatttgataagaaagatgcttaaagtgctactagaacaagagaaa 97
            |||||| ||  |||||||| |||||||||||||||| || |||| |||||||| ||
Db      77 CAGGAATAAACACGTTTGGTGAGAGCCCATGGCACTCAAGGTCCTACCTCHACAGGACG 136

Qy      98 acgtttttcactctttagtattactaggctcattgtcatgtaaagtgacttgtgaatca 157
            | |||  |  ||||| || | || ||| ||| |   ||||| ||||| ||| | |||
Db     137 GTGCTCTTCGGTCGCCAHTCTCTTCCTACTCCACCTGGCCATGTAAAGTGAGTTGCGAAACC 196

Qy     158 ggagactgtagaacagcaagaattcaggagatcggtctggaaactgtgttccctgcaaccag 217
            |||| |||  || || |  |  |||||||||||| || ||||||||||||| |||||
Db     197 GGAGAATGCAGGCAGGAATTCAAGGATGCGATCTGGAAACTGTGTTCCTCTGCAAACAG 256

Qy     218 tgtgggccaggccaatggaagtgtctaaggaatgtggcttccggcgctatggggagggg 277
            || |||||||| |||||||| ||||| | |||||||| ||||||||||||||||| ||
Db     257 TGCGGACCTGGCCATGGAAGTTGTCCAAGGAATGTGGCTTCGGCTAGGGGGAGGGGATGCACAG 316

Qy     278 tgtgtgggctgtccggcgtgcacagagttcaagggagggcttcaagaaatgtcaagccc 337
            |||| ||||| |||  || ||||||| ||||||||| ||||| |||||| || || ||||
Db     317 TGTGTGCCCTGCAGGCGCCGCACACGGTTCAAGGGAGTGTAAGCCCA 376

Qy     338 tgtctggactgcgcagtggtgaaccgcgctctcagaaggcaaattgttcagccaccagtgat 397
            |||| |||||| |||||||||| || ||| ||||||||| ||||||||||||||||||
Db     377 TGTGCCGGACTGCGCTGGTGAACCCCTTTCAGAGGGCCCACTGCGTCACACACCAGTGAT 436

Qy     398 gccatctgcggggactgcgttgcccaggacggattttatatagaagaagaaactttgtgcttca 457
            || || ||||||||||||||| ||||||||||| ||||||||| || |||||||||||
Db     437 GCTGTCTGCGGGGACTGCGTGCCTCCAGGAGTTTACCGAAGACCAAACTGGTGTTGGCTTCA 496

Qy     458 gacatggagtgtgtgcttgtggagaccctcctcctctcctacgaaccgcactgtgtccagc 517
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     497 GACATGGAGTGTGTGCCCTGCGGGGAGACCACTCTCCTCCTACGAACCACACTGTGAGTGA 556

Qy     518 aaggtcaacctcgtgaagatc 538
            | |||
```

```
Db                   557 TGTGCCAAGTGGCAGAGACC 577

RESULT 11
X23416
ID   X23416 standard; DNA; 1133 BP.
AC   X23416;
DT   18-JUN-1999 (first entry)
DE   Rat rAPO4-alpha DNA.
FH   Key              Location/Qualifiers
FT   CDS              1..330
FT                    /*tag= a
FT                    /product= "rAPO4-alpha"
PN   WO9911791-A2.
PD   11-MAR-1999.
PF   04-SEP-1998; U18393.
PF   05-SEP-1997; US-924634.
PA   (UNIW ) UNIV WASHINGTON.
PI   Chaudhary PM;
DR   WPI; 99-205191/17.
DR   P-PSDB; W95582.
PT   New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT   useful for diagnosis and treatment of prostate cancer and
PT   developmental or gestational abnormalities
PS   Example IV; Fig 7D; 156pp; English.
CC   This invention describes isolated Tumor Necrosis Factor (TNF) family
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC   their active fragments. APO4 is useful for diagnosing prostate cancer
CC   by determining levels of APO4 in an individual. Prostate cancer can also
CC   be treated using APO4 binding agents linked to a therapeutic
CC   moiety. APO4 polypeptides are also useful for identifying selective
CC   binding agents, useful in diagnosis/treatment of disease by binding of
CC   agents to the polypeptide/active fragment which is extracellular, or
CC   expressed on the cell surface. The binding is preferably performed in
CC   vivo. APO4 polypeptides/active fragments are also useful for screening
CC   for agonists and antagonists by binding and observing the change in APO4
CC   activity. Effective pharmacological agents useful in diagnosis or
CC   treatment of disease are also identified using APO4 polypeptides/active
CC   fragments and APO4 signal transducer molecules that specifically interact
CC   with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC   activity. The method is performed in vivo or in vitro. Apo polypeptides
CC   are all useful as immunogens for preparing antibodies. APO4 is also
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma cell line
CC   MCF-7, and induced apoptosis.
SQ   Sequence 1133 BP;      278 A;      275 C;      258 G;     296 T;

Qy   Query Match      8.1%;  Score 120.2;  DB 1;  Length 1133;
     Best Local Similarity   62.8%;
     Matches 204;  Conservative   0;  Mismatches 118;  Indels   3;  Gaps   1;

Qy     985 tggcctctgatgcagaatcccatgggttggtgacaacatctctttgtgactcttatcct 1045
            ||||| ||||| | |||||||||||||||||||||| || ||||||||| || || ||
Db       1 TGGCCTCTGATGCAGATATCCCTTCTGGGGTGGTG---ACAGCTCTCTCTGTACTCTTTATCCT 57

Qy    1046 gaactcactggagaagaacaattcattctctcaatccagaaacttgaaagctcaacgtcttg 1105
            ||||| || ||| ||||| |||| ||| ||||||||| |||||||||||||||||| |||
Db      58 GAACTCACTGGAGAAGATATCAATTCCCTCAATCCCAGAAAATGAAAGCTCAACATCTGTG 117

Qy    1106 gattcaaatagacagtcaagaattttggttggtgtgggggctgttccagtccagtctcattcgaa 1165
            ||| |||| | | |||||||| ||||||||| ||| ||||||||| ||| ||||||||||| |
Db     118 GATTCCAAATGGTGGTCAGGATTTGGGGCTGGGGCAGAGCTGGGCTGCTTCCAGATTTCTTCTGAGATTTT 177

Qy    1166 aacttacagcagctactgattttatctagatataacaacacactggtagaacagcagcatca 1225
            | || | || | ||  |  |  ||  ||  ||  |  | |  |  | |  ||   ||  |
```

```
Db    178 CAAGAAAATACTGACTTCACCTACGATATATTGGTGACGCGGCATACGGTCTGGGGACGCAAAACCTA 237
          |||||||||||||| |||||||||||||||||||||| |||||||||||||||
Qy   1226 actcaggatgtcactactatagagaccagctagatcaggagagtggcggctgtcatccac 1285
Db    238 GCTCAGGATGTCTCAAAGGACTCCAAGCACGAGAGGGCTGGGAAGCCAGTGAAAACCTGAAT 297
          ||||| ||||||| |||||||||||||
Qy   1286 ccagccactcaagacgtcccctccagg 1310
          |||||| ||||||||||||||||
Db    298 CTAGCCACGTCCACAGCCTTCCAGG 322
```

RESULT 12

```
ID   X24979 standard; cDNA; 791 BP.
AC   X24979;
DT   05-JUL-1999 (first entry)
DE   Clone GJ156 encoding TRAIN-R secreted form C-terminus.
KW   TRAIN-R; receptor; human; tumour necrosis factor receptor;
KW   agonist; antagonist; cancer; immunological disease; therapy;
KW   cytostatic; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Homo sapiens.
FT   intron          1..350
FT                   /*tag= a
FT   exon            351..790
FT                   /*tag= b
FT   CDS             352..444
FT                   /*tag= c
FT                   /partial
FT   3'UTR           45..790
FT                   /product= "TRAIN-R secreted form C-terminus"
FT                   /*tag= d
```

```
PN   WO9913078-A1.
PD   18-MAR-1999.
PF   11-SEP-1998; U19030.
PR   06-MAY-1998; US-084422.
PR   12-SEP-1997; US-058631.
PA   (BIOJ ) BIOGEN INC.
PI   Hession C, Tschopp J;
DR   WPI; 99-222238/19.
DR   P-PSDB; W98147.
PT   New cysteine-rich tumor necrosis factor receptor
PS   Claim 1; Page 28; 30pp; English.
CC   The present sequence includes an exon encoding the C-terminus (see
CC   W98147) of a soluble form of a novel human cysteine-rich tumour
CC   necrosis factor receptor family member termed TRAIN-R. It comprises
CC   clone GJ156, obtained from a Clontech human adult lung cDNA library.
CC   The encoded 30-amino acid C-terminal peptide is identical to amino
CC   acids 121-149 of the composite TRAIN-R protein given in W98146 and
CC   to amino acids 121-150 of the composite TRAIN-R of murine TRAIN-R short
CC   form (secreted protein, see W98144). The soluble protein is
CC   expected to inhibit signaling by the full-length TRAIN-R. Human
CC   TRAIN-R is expressed at low levels in every tissue and cell line
CC   tested thus far, with higher expression detected in heart, prostate,
CC   ovary, testis, peripheral blood lymphocytes, thyroid and adrenal
CC   gland. Cell death can be induced by administering an agent capable
CC   of inhibiting the binding of TRAIN-R to its ligand. A claimed method
CC   of treating or reducing, the advancement, severity or effects of
CC   an immunological disease in a mammal comprises administering a
CC   pharmaceutical composition which comprises a TRAIN-R blocking agent,
CC   e.g. soluble TRAIN-R. TRAIN-R can be fused to an immunoglobulin to
CC   produce a fusion protein which may be targeted to various sites.
CC   It can be used in binding assays, and to identify antagonists and
CC   agonists. Anti-TRAIN-R antibodies can be used to reduce the
CC   severity of an immune response or to treat cancer. TRAIN-R
CC   blocking agents can also be used to reduce the severity or effects
CC   of an immunological disease (all claimed).
SQ   Sequence 791 BP; 202 A; 189 C; 165 G; 235 T;
```

```
Query Match             6.1%;  Score 90.4; DB 1; Length 791;
Best Local Similarity   89.8%; Pred. No. 3.6e-16;
Matches    97; Conservative   0; Mismatches  11; Indels   0; Gaps   0;
```

```
Qy    413 tgcttgccaggatttatagaggaagaccgaaaacttgtcggcttcaagacgatggagtgtgtg 472
          ||||| |||||||||| |||||||||||||||||||||| |||||||
Db    340 TTCTCTTCTAGATTTATAGGACGAAAACCTTGTCGGCTTTCAAGACACATGGAGAGTGTGTG 399
Qy    473 ccttgtgggaggaccgtccctcctccttacgaaccgcactgtgccaagcaag 520
          ||||||||||||||||||||||||||||||||||||||||||| |||
Db    400 CCTTGTGGGAGGACCCTCCCTCCTCCTTACGAACCGCACTGTGACTGAGCAAG 447
```

RESULT 13

```
ID   V11423 standard; cDNA; 181 BP.
AC   V11423;
DT   23-JUL-1998 (first entry)
DE   Human secreted protein clone AX92_3 cDNA internal fragment.
KW   Secreted protein; prevention; treatment; gene therapy; ds.
OS   Homo sapiens.
DR   WPI; 99-110230/10.
DR   P-PSDB; W58844.
PT   Secreted proteins and polynucleotides encoding them - useful to
PT   prevent, treat and ameliorate medical conditions
PS   Claim 13; Page 57; 93pp; English.
PN   WO9801554-A2.
PD   15-JAN-1998.
PF   07-JUL-1997; U11876.
PR   09-JUL-1996; US-677231.
PA   (GEMY ) GENETICS INST INC.
PI   Bowman M, Evans C, Jacobs K, LaValle ER, McCoy JM,
PI   Merberg D, Racie LA, Spaulding V, Treacy M;
CC   V11422-V11424 encode fragments of a novel secreted protein derived from
CC   clone AX92_3 which was isolated from a human fetal brain cDNA library.
CC   The protein can be used to prevent, treat or ameliorate a medical
CC   condition, while the polynucleotides can be used for gene therapy.
SQ   Sequence 181 BP; 43 A; 47 C; 34 G; 57 T;
```

```
Query Match             5.9%;  Score 87.4; DB 1; Length 181;
Best Local Similarity   98.9%; Pred. No. 1.1e-15;
Matches    88; Conservative   0; Mismatches  1; Indels   0; Gaps   0;
```

```
Qy    423 gatttatagaggaagacgaaacttgtcggcttcaagacgacatggagtgtgtgccttgtgggag 482
          |||||||||||||||||||||| |||||||| |||||||||||||||||||||||||||||||||
Db     88 GATTTTATAGGACGAAAACCTTGTCGGCTTTCAAGACACATGGAGTGTGTGCCTTGTGGGAG 147
Qy    483 accctccctcctccttacgaaccgcactgt 511
          ||||||||||||||||||||||||||||||
Db    148 ACCCTCCCTCCTCCTTACGAACCGCACTCT 176
```

RESULT 14

```
ID   V86655 standard; cDNA; 201 BP.
AC   V86655;
DT   27-APR-1999 (first entry)
DE   EST clone AX92.
KW   Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW   tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW   receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
KW   chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
OS   Homo sapiens.
PN   WO9845435-A2.
PD   15-OCT-1998.
PF   10-APR-1998; U06954.
PR   10-APR-1997; US-835913.
PA   (GEMY ) GENETICS INST INC.
PI   Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
PI   Racie LA, Spaulding V, Treacy M;
DR   WPI; 99-070076/06.
PT   New polynucleotides encoding human secreted proteins - derived from
PT   e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT   ovary, pituitary, retina and colon cDNA libraries
PS   Claim 1; Page 306; 633pp; English.
```

CC   This sequence represents an expressed sequence tag (EST), and is a
CC   polynucleotide of the invention. The polynucleotides of the invention are
CC   all secreted EST sequences isolated from a variety of human tissue
CC   sources. The EST sequences and proteins encoded by them are predicted to
CC   have useful biological activities which would make them suitable for
CC   treating, preventing or ameliorating medical conditions in humans and
CC   animals, although no supporting data is given. Suggested activities
CC   include nutritional activity, immune stimulating or suppressing activity,
CC   haematopoiesis regulating activity, tissue growth activity,
CC   activity/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC   activity. The EST sequences are also stated to be useful for gene
CC   therapy.
SQ   Sequence   201 BP;   52 A;   51 C;   38 G;   60 T;

Query Match   5.9%; Score 87.4; DB 1; Length 201;
Best Local Similarity   98.9%; Pred. No. 1.2e-15;
Matches   88; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   423 gatttattaggaagacgaaacttgtccggcttcaagacgacactgggagtgtgtgcctgggag 482
          ||||||||||||||||||||||||||| || |||||||| |||| ||||||||||||||| |||
Db   93 GATTTATTAGGAAGACGAAACTTGTCGGCGTTCAAGACAACATGGAGTGTGTGCCTGTGGAG 152

Qy   483 accctccctcccttacgaaccgcactgt 511
          ||||||||||||||||||||||||| ||
Db   153 ACCCTCCCTCCCTTACGAACCGCACTCT 181

RESULT 15
X23418
ID   X23418 standard; DNA; 396 BP.
AC   X23418;
DT   18-JUN-1999 (first entry)
DE   Mouse mAPO4-beta DNA.
KW   Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW   developmental abnormality; gestational abnormality; prostate cancer;
KW   APO6; APO8; APO9; TNRL1; TNRL-3; diagnosis; treatment; therapy; disease;
KW   cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW   apoptosis; mouse; APO4-beta; ss.
OS   Mus sp.
FH   Key   Location/Qualifiers
FT   CDS   1..289
FT   /tag= a
FT   /product= "mAPO4-beta"
PN   WO9911791-A2.
PD   11-MAR-1999;
PF   04-SEP-1998; U18393.
PR   05-SEP-1997; US-924634.
PA   (UNIW) UNIV WASHINGTON.
PI   Chaudhary PM;
PI   WPI; 99-205191/17.
DR   P-PSDB; W93584.
PT   New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT   useful for diagnosis and treatment of prostate cancer and
PT   developmental or gestational abnormalities
PS   Example IV; Fig 8; 156pp; English.
CC   This invention describes isolated Tumor Necrosis Factor (TNF) family
CC   receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC   fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC   their active fragments. APO4 is useful for diagnosing prostate cancer
CC   by determining levels of APO4 in an individual. Prostate cancer can also
CC   be treated using APO4 selective binding agents linked to a therapeutic
CC   moiety. APO4 polypeptides are also useful for identifying selective
CC   binding agents, useful in diagnosis/treatment of disease by binding of
CC   agents to the polypeptide/active fragment which is extracellular, or
CC   expressed on the cell surface. The binding is preferably performed in
CC   vivo. APO4 polypeptides/active fragments are also useful for screening
CC   for agonists and antagonists by binding and observing the changer in APO4
CC   activity. Effective pharmacological agents useful in diagnosis or
CC   treatment of disease are also identified using APO4 polypeptides/active
CC   fragments and APO4 signal transducer molecules that specifically interact

CC   with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC   activity. The method is performed in vivo or in vitro. APO polypeptides
CC   are all useful as immunogens for preparing antibodies. APO4 is also
CC   useful for diagnosis/treatment of developmental or gestational
CC   abnormalities. APO8 was transfected to human breast carcinoma cell line
CC   MCF-7, and induced apoptosis.
SQ   Sequence   396 BP;   107 A;   102 C;   108 G;   79 T;

Query Match   5.4%; Score 80.4; DB 1; Length 396;
Best Local Similarity   55.8%; Pred. No. 1.7e-13;
Matches   153; Conservative   0; Mismatches   121; Indels   0; Gaps   0;

Qy   1037 tcttatcctgaactcactggagaaggacattcaattcctcaatccagaaacttgaaagtct 1096
          || |||||||||| |||||||||||| || | ||||||| |||| ||||||| |||| ||
Db   8 TCCGATCCTGAAACTCACTGGAGAAGATAACCAAATTCCCTCAANTCCCGAAAAACGCA 67

Qy   1097 acgtgtcttttgyattcaaatagcagtcaagatttggttggtggggggtcgtgttccagtccagtct 1156
          ||| | | | ||||| | || ||| | ||| |||| ||||||||| | || |||||| | |
Db   68 GCATCTCTTGGATTCAGTTGCGGGCCCAGGATCTGGCGTTGGGGACAGCTGCTCTTAGAGTGTCTT 127

Qy   1157 cattctgaaacttacagacgctactgatttatctagatatataacaacacactggttagaa 1216
          || ||| ||| ||||| | |||| |||| |||||||| |||||||||||||||| |||||
Db   128 GGGAANTGTTTCAGAANTCTACTGACTCACCTAGACACATGGTGACACTGGTACAGTCTGGGAG 187

Qy   1217 tcagcatcaactcaggatgtcactaactatagaagaagccagtctagatcaaggagagagtggcgct 1276
          |||| | | |||||||||| | | || |||||||||||| || |||||||||| || |||
Db   188 CAGAGCGTACGTCAGGATGTGCTTCAAAGGAGCCAAGGAGGCTGGAAGGAGGACAGGGAA 247

Qy   1277 gtcatccaccagcccactcaggacgtccctccagg 1310
          | | | ||| ||| ||||| | ||| ||||||
Db   248 AACCTGAAATCTAGCCATGCCCACAGCAGCCTTCCAGG 281

Search completed: April 23, 2000, 02:02:33
Job time: 1272 sec

OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 23, 2000, 01:28:55 ; Search time 700.48 Seconds
(without alignments)
-6454.397 Million cell updates/sec

Title: US-09-490-187-1
Perfect score: 1489
Sequence: 1 ggacctgcagcctcccaggt.........gtattttttaaaaacttt 1489

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 821193 seqs, -1518192014 residues

Total number of hits satisfying chosen parameters: 1642386

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*

```
 1: gb_bal:*        24: em_pat:*
 2: gb_ba2:*        25: em_ph:*
 3: gb_om:*         26: em_pl:*
 4: gb_ov:*         27: em_ro:*
 5: gb_pat:*        28: em_sts:*
 6: gb_ph:*         29: em_sy:*
 7: gb_pl1:*        30: em_un:*
 8: gb_pl2:*        31: em_vi:*
 9: gb_pr1:*        32: gb_htg1:*
10: gb_pr2:*        33: gb_htg2:*
11: gb_pr3:*        34: gb_in1:*
12: gb_ro:*         35: gb_in2:*
13: gb_sts:*        36: em_bal:*
14: gb_sy:*         37: em_ba2:*
15: gb_un:*         38: em_hum3:*
16: gb_vi:*         39: em_hum4:*
17: em_fun:*        40: gb_pr4:*
18: em_hum1:*       41: gb_htg3:*
19: em_hum2:*       42: gb_htg4:*
20: em_in:*         43: gb_htg5:*
21: em_om:*         44: gb_htg6:*
22: em_or:*         45: gb_htg7:*
23: em_ov:*         46: em_htg1:*
                    47: em_htg2:*
                    48: em_htg3:*
                    49: em_hum5:*
                    50: gb_pl3:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	41.2	2.8	26258	34	CEL2K816	U41018 Caenorhabdi
2	40.8	2.7	123416	33	AC004847	AC004847 Homo sapi
3	40.6	2.7	149	5	A08919	A08919 H.saplens (
4	40.4	2.7	207699	42	AC012074	AC012074 Homo sapi
5	40.4	2.7	215566	42	AC012365	AC012365 Homo sapi
6	40.2	2.7	1111944	32	HSDJ130E4	AL078582 Homo sapi
7	40.2	2.7	237316	42	AC011599	AC011599 Homo sapi
8	40	2.7	3417	7	LEGGFRPOL	X80908 L.esculentu
9	40	2.7	173767	10	AC002449	AC002449 Human PAC
10	39.8	2.7	2670	7	D88490	D88490 Candida alb
11	39.6	2.7	4673	34	PFASSA	M84732 Plasmodium
12	39.6	2.7	299820	33	AC006871	AC006871 Caenorhab
13	39.6	2.7	299820	33	AC006803	AC006803 Caenorhab
14	39.4	2.6	1840	34	DDU73686	U73686 Dictyostel
15	39.4	2.6	80331	40	AC004994	AC004994 Homo sapi
16	39.4	2.6	98864	11	HS460D19	AL031905 Human DNA
17	39.4	2.6	110000	32	CEY73F8_2	Continuation (3 of
18	39.4	2.6	216625	34	AC012671	AC012671 Homo sapi
19	39.4	2.6	322774	34	CEY73F8A	AL132862 Caenorhab
20	39.2	2.6	720	8	CNS01BX6	CNS01BX6 Botrytis
21	39.2	2.6	38532	1	SCE94	AL114802 Streptomy
22	39.2	2.6	184811	43	AC009399	AC009399 Homo sapi
23	39.2	2.6	321003	32	PFMAL4P3	AL034476 Plasmodiu
24	39	2.6	8586	43	AC015290	AC015290 Drosophil
25	38.8	2.6	2271	10	HSM800417	HSM800417 Homo sapi
26	38.8	2.6	1582	5	A77033	A77033 Sequence 10
27	38.8	2.6	1582	5	A77035	A77035 Sequence 12
28	38.6	2.6	114467	10	AP000462	AP000462 Homo sapi
29	38.6	2.6	129537	33	AC007775	AC007775 Homo sapi
30	38.6	2.6	157825	10	CNS01DRI	CNS01DRI Human chr
31	38.4	2.6	196287	10	CNS0000B	CNS0000B Human chr
32	38.4	2.6	200509	33	AC006910	AC006910 Caenorhab
33	38.4	2.6	207370	33	AC006798	AC006798 Caenorhab
34	38.2	2.6	163612	43	AC012213	AC012213 Homo sapi
35	38.2	2.6	178972	44	AC016369	AC016369 Homo sapi
36	38.2	2.6	182374	44	AC008114	AC008114 Homo sapi
37	38.2	2.6	204652	32	PFMAL13P6	AL049183 Plasmodiu
38	38	2.6	14433	35	AE001369	AE001369 Plasmodiu
39	38	2.6	58516	42	AC011982	AC011982 Homo sapi
40	38	2.6	173648	43	AC015861	AC015861 Homo sapi
41	38	2.6	196149	33	AC004709	AC004709 Homo sapi
42	38	2.6	208807	41	AC004688	AC004688 Plasmodiu
43	38	2.6	259474	40	HUAC004605	AC004605 Plasmodiu
44	37.8	2.6	31950	34	CEC38H2	Z56641 Caenorhabdi
45	37.8	2.5	52238	33	AC007972	AC007972 Homo sapi

ALIGNMENTS

RESULT 1
CEL2K816/c

LOCUS   CEL2K816   26258 bp   DNA   INV   30-NOV-1995
DEFINITION Caenorhabditis elegans cosmid 2K816.
ACCESSION   U41018
VERSION   U41018.1   GI:1086666
KEYWORDS

```
SOURCE
  ORGANISM  Caenorhabditis elegans strain=Bristol N2.
            Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
            Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 26258)
  AUTHORS   Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M.,
            Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J.,
            Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A.,
            Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M.,
            Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N.,
            Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B.,
            O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A.,
            Saunders,D., Shownkeen,R., Smaldon,N., Smith,A., Sonnhammer,E.,
            Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M.,
            Vaughan,K., Waterston,R., Watson,A., Weinstock,L.,
            Wilkinson-Sproat,J. and Wohldman,P.
  TITLE     2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
            elegans
  JOURNAL   Nature 368 (6466), 32-38 (1994)
  MEDLINE   94150718
REFERENCE   2  (bases 1 to 26258)
  AUTHORS   Nhan,M and Le,T.
  TITLE     The sequence of C. elegans cosmid ZK816
  JOURNAL   Unpublished (1995)
REFERENCE   3  (bases 1 to 26258)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-1995) Robert Waterston
COMMENT     Genome Sequencing Center
            Department of Genetics, Washington University,
            St. Louis, MO 63110, USA, and
            Sanger Centre, Hinxton Hall
            Cambridge CB10 1RQ, England
            e-mail: rw@nematode.wustl.edu and jes@sanger.ac.uk
            NEIGHBORING COSMID INFORMATION:
The 5' cosmid is F11D5, 200 bp overlap; 3' cosmid is ZK563, 200 bp
overlap. Actual start of this cosmid is at base position 35525 of
CELF11D5; actual end is at 6673 of CELZK563.

NOTES:
Coding sequences below are predicted from computer analysis, using
the program Genefinder(P. Green and L. Hiller, ms in preparation).
FEATURES             Location/Qualifiers
     source          1..26258
                     /organism="Caenorhabditis elegans"
                     /db_xref="taxon:6239"
     gene            3410..5744
                     /gene="ZK816.2"
                     /map="X"
     CDS             join(3410..3500,3603..3943,4004..4066,4112..4198,
                     5004..5219,5269..5485,5624..5681,5729..5744)
                     /gene="ZK816.2"
                     /note="coded for by C. elegans cDNA yk5g6.5; coded for by
                     C. elegans cDNA yk123g1.5; coded for by C. elegans cDNA
                     yk123g1.3; coded for by C. elegans cDNA yk5g6.3; similar
                     to S. cerevisiae phosphate system positive regulatory
                     protein PHO4 (SP:PHO4_YEAST, P07270)"
                     /codon_start=1
                     /protein_id="AAA82323.1"
                     /db_xref="GI:1086667"
                     /translation="MISRYTNPPPTSGTGCSSFQMRWVSSALRTYDKTDETTTDSMRS
                     TAPLRPPITYSPSYTSTSLNSVSLSSKYVEEKSKTSQGSTAPVVPSNSLSGLLSPREK
                     PYGDDVTILKSVSERVPENFELRRKQQVKALFRIHGRCLKFAEPTNETQKLAEPSATN
                     IVNQQDSDDNSDGDSRDGNSDYIDIFDMRPNSLVFTRLANEEEMGDGDNVISRSLLSL
                     VAEEDLPDFIVALNEKPIRYVWVEESEKPTSSSILRSRDARKEEKVQKGSTRIRFCTD
                     ANDNIHPLTHDLTFSDSCLETSAVSSAY"
     gene            7063..7361
                     /gene="ZK816.3"
     CDS             join(7063..7064,7110..7175,7229..7361)
                     /gene="ZK816.3"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAA82324.1"
                     /db_xref="GI:1086668"
                     /translation="MNASQQRPLPFFFRLFIGPHIARYLRMRRLMNEARRQNLMAHRR
                     QDNANTSSNSSVVQRSSVSETY"
     gene            complement(8331..9120)
     CDS             complement(join(8331..8479,8533..8660,8711..8827,
                     8878..8954,9001..9120))
                     /codon_start=1
                     /gene="ZK816.4"
                     /evidence=not_experimental
                     /protein_id="AAA82325.1"
                     /db_xref="GI:1086669"
                     /translation="MLAKNCILMLISLSCYCQFKFPPPTQEFFLDRSNGYDYNSSEN
                     FERPLNASMYQPTTCGEVEQERVKECADPLYKMGATSENSHYLGWGGFIFRTKAFSE
                     VCDNFLFDVCTEPYRKDVCFAADRARFNYDAAIKILDFLCRDGYGEMLRNIECFTKTL
                     TRSEMMOCQAELVSDTPRKISESHSEVKGANDAAVCG"
     gene            19067
                     /gene="ZK816.1"
     CDS             join(17796..17890,18804..18878,18923..19067)
                     /gene="ZK816.1"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAA82326.1"
                     /db_xref="GI:1086670"
                     /translation="MVLVILCFVLPAIIFFVEALIAWLRVLQLNYRRRRATEAADREA
                     KKKASQLESLITKDGKVICHIPIQVEDIETGAMFIYNSDENLAIIGEDAKIPLVEQO
                     LV"
     gene            complement(19490..21064)
                     /gene="ZK816.5"
     CDS             complement(join(19490..19629,20019..20172,20221..20395,
                     20447..20522,20699..20961,20995..21064))
                     /gene="ZK816.5"
                     /note="coded for by C. elegans cDNA yk102d10.5; similar
                     to short-chain alcohol dehydrogenase."
                     /codon_start=1
                     /protein_id="AAA82327.1"
                     /db_xref="GI:1086671"
                     /translation="MSLKSKIAIVIGASRGGRGMLKGVALQLAEAGCTLYIIGRAPS
                     KTLSSELTYLFPTLEGTAEEECRKRGGICHVRVDHSNMDEVKFFDEVASEFDNQLDII
                     VNNASAVTKCGSGDVRKFEEMDGGVNEEIWDIINNVGLRNQYCSVVGTRIMRKNGMKGLI
                     VNISSLGGIMYLFTVAYGVGKMALDRMSSDMAQELQDIGITIVISMPSAVKTELITNM
                     IETSAGSWGAYENKMFLNGESTEYCGKAVVAIAADPKKRKVNGSLITTLDMGNYSYT
                     DIDGRLPTNMRQLRGLLSLAGYHSMAGWCPEWVNLPGWAIILWQNKIHY"
BASE COUNT     8375 a   4735 c   4187 g   8961 t
ORIGIN

RESULT 2

Query Match     2.8%;  Score 41.2;  DB 34;  Length 26258;
Best Local Similarity   67.4%;
Matches  58;  Conservative   0;  Mismatches  28;  Indels   0;  Gaps   0;

Qy   1400  atgagtttttttttttttgcatcttttaataatttcctatatgttgtagagtatgtttaaaa  1459
           | |||||  |||||||||||||||||||||||| ||||| ||||| ||| |||||||
Db   4311  ATTAGTTATTTTTTTTGGATGTTATTTTTTTTTGTTACTATGTTCCAAATTATATTTTTATATA  4252

Qy   1460  taaattcaagtattttttttaaaaa  1485
           |||||||| || ||| |||||||
Db   4251  TGAATTTAACATTTCATTTTACAAA  4226

RESULT 2
AC004847/c
LOCUS       AC004847    123416 bp    DNA    HTG    12-JUN-1998
DEFINITION  Homo sapiens clone DJ0647J21, *** SEQUENCING IN PROGRESS ***, 10
            unordered pieces.
ACCESSION   AC004847
VERSION     AC004847.1  GI:3213150
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
```

```
REFERENCE   1  (bases 1 to 123416)
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS     Waterston,R.H.
TITLE       The sequence of Homo sapiens clone
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 123416)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 10 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *       1    1580: contig of 1580 bp in length
          *    1581    1598: gap of unknown length
          *    1599    3512: contig of 1914 bp in length
          *    3513    3530: gap of unknown length
          *    3531    6664: contig of 3134 bp in length
          *    6665    6682: gap of unknown length
          *    6683   11088: contig of 4406 bp in length
          *   11089   11106: gap of unknown length
          *   11107   20349: contig of 9243 bp in length
          *   20350   20367: gap of unknown length
          *   20368   33197: contig of 12830 bp in length
          *   33198   33215: gap of unknown length
          *   33216   43072: contig of 9857 bp in length
          *   43073   43090: gap of unknown length
          *   43091   65746: contig of 22656 bp in length
          *   65747   65764: gap of unknown length
          *   65765   85268: contig of 19504 bp in length
          *   85269   85286: gap of unknown length
          *   85287  123416: contig of 38130 bp in length
FEATURES             Location/Qualifiers
     source          1..123416
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DJ0647J21"
BASE COUNT    30903 a  29142 c  29866 g  33343 t    162 others
ORIGIN

RESULT   3
A08919/c
LOCUS       A08919     149 bp     DNA         PAT    02-SEP-1993
DEFINITION  H.sapiens (haplotype 2A, allele MS32, isolate English, serial
            number 30) minisatellite sequence.
ACCESSION   A08919
VERSION     A08919.1  GI:411841
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 149)

Query Match           2.7%;  Score 40.8;  DB 33;  Length 123416;
Best Local Similarity 59.5%;  Pred. No. 14;  Mismatches 47;  Indels 0;  Gaps 0;
Matches 69;  Conservative 0;

Qy     1365 acctctgttcccagccataagatttgggggaacctggatcgagtttttttttgcatcctta 1424
            |||||||| |||  || | ||||||||| || ||||||||||||  |||
Db     34675 ACTCTGTTCAGATGGCAGGACATGAGAGTCAAAGAAGAGATTGTTCTTCTTTTTTTAATTTTA 34616

Qy     1425 ataatttccttatatgttgtagagtatgttttaaaataaatttcaagtattttttta 1480
            | ||| || | || | ||  ||||| | ||| |||||| ||||||||| ||||||||
Db     34615 AAAAATATTTTAAATATATACATTATTATTTTTTACTTAAAAATTTACGTATTTATTTA 34560
```

```
AUTHORS     Jeffreys,A.J.
TITLE       Extended nucleotide sequences
JOURNAL     Patent: EP 0370119-A 74 30-MAY-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES             Location/Qualifiers
     source          1..149
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      100 a      0 c      0 g     49 t
ORIGIN

Query Match           2.7%;  Score 40.6;  DB 5;  Length 149;
Best Local Similarity 66.7%;  Pred. No. 6.9;
Matches 58;  Conservative 0;  Mismatches 29;  Indels 0;  Gaps 0;

Qy     1403 agtttttttttgcatccttaataattccttatatgttgtagagtatgttttaaaataa 1462
            |||||||||||| | |||||||||| |||||| || ||| | ||| ||  || || |
Db       92 ATTTTTTTTTTTTTTTTTTTTTTTTAATTTTTTTTTTTTTTATTTTTTATTTTTTTTATA 33

Qy     1463 atttcaagtatttttttaaaaaactt 1489
            |  |||||||| |||||| ||| |
Db       32 ATTTTATTTTTTTTTTTTTTTAATTAAAAATAT 6
```

```
RESULT   4
AC012074
LOCUS       AC012074   207699 bp   DNA         HTG    19-OCT-1999
DEFINITION  Homo sapiens clone NH0458N05, *** SEQUENCING IN PROGRESS ***, 40
            unordered pieces.
ACCESSION   AC012074
VERSION     AC012074.1  GI:6067201
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 207699)
AUTHORS     Waterston,R.H.
TITLE       The sequence of Homo sapiens clone
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 207699)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (19-OCT-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 40 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *       1    2641: contig of 2641 bp in length
          *    2642    2660: gap of unknown length
          *    2661    5448: contig of 2788 bp in length
          *    5449    5467: gap of unknown length
          *    5468    9447: contig of 3980 bp in length
          *    9448    9466: gap of unknown length
          *    9467   13258: contig of 3792 bp in length
          *   13259   13277: gap of unknown length
          *   13278   17777: contig of 4500 bp in length
          *   17778   17796: gap of unknown length
          *   17797   22203: contig of 4407 bp in length
          *   22204   22222: gap of unknown length
          *   22223   25932: contig of 3710 bp in length
          *   25933   25951: gap of unknown length
          *   25952   29791: contig of 3840 bp in length
          *   29792   29810: gap of unknown length
          *   29811   33884: contig of 4074 bp in length
          *   33885   33903: gap of unknown length
```

```
Qy    1395 cctggatgagtgtttttttttttgtgtcatctttaataatttctttatatgtgttgaagtatgtgtt 1454
               ||||||||||||||||| |||||||||||||||||||| |||||||||| |||| ||||| ||| |
Db  127387 CGTGGCTTATTTATTTATTATTTTTTAAACTATPAATCTTCTGGAATTTGCTTTAGAATATTATG 127446

Qy    1455 taaaataaattcaagtgtattttttaaaa 1484
               |||||||||||||| |||||||| ||||
Db  127447 TAAAAATAAATGTAAAATTATTATTTTTATA 127476

RESULT  5
LOCUS       AC012365       215566 bp       DNA                      HTG       25-OCT-1999
DEFINITION  Homo sapiens clone NH0459C22, *** SEQUENCING IN PROGRESS ***, 37
            unordered pieces.
ACCESSION   AC012365
VERSION     AC012365.1  GI:6114946
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 215566)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 215566)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-OCT-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 37 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *         1:  contig of 6252 bp in length
            *      6253:  gap of unknown length
            *      6271:  contig of 7045 bp in length
            *     13316:  gap of unknown length
            *     13317:  contig of 2853 bp in length
            *     13335:  gap of unknown length
            *     13336:  contig of 7894 bp in length
            *     21229:  gap of unknown length
            *     21230:  contig of 6346 bp in length
            *     27594:  gap of unknown length
            *     27595:  contig of 8359 bp in length
            *     35972:  gap of unknown length
            *     35973:  contig of 9680 bp in length
            *     45671:  gap of unknown length
            *     45672:  contig of 13020 bp in length
            *     58710:  gap of unknown length
            *     58730:  contig of 12648 bp in length
            *     71377:  gap of unknown length
            *     71396:  contig of 18612 bp in length
            *     90008:  gap of unknown length
            *     90009:  contig of 31565 bp in length
            *    121592:  gap of unknown length
            *    121610:  contig of 2470 bp in length
            *    124080:  gap of unknown length
            *    124081:  contig of 2262 bp in length
            *    126360:  gap of unknown length
            *    126361:  contig of 2853 bp in length
            *    129231:  gap of unknown length
            *    129250:  contig of 2163 bp in length
            *    131412:  gap of unknown length
            *    131430:  contig of 2605 bp in length
            *    134035:  gap of unknown length
            *    134053:  contig of 2138 bp in length
            *    136191:  gap of unknown length
            *    136209:  contig of 2395 bp in length
```

```
FEATURES             Location/Qualifiers
     source          1..207699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NH0458N05"
     *      33904:  contig of 4268 bp in length
     *      38171:  gap of unknown length
     *      38172:  contig of 18 bp in length
     *      38190:  gap of unknown length
     *      38191:  contig of 5547 bp in length
     *      43737:  gap of unknown length
     *      43738:  contig of 18 bp in length
     *      43755:  gap of unknown length
     *      43756:  contig of 5222 bp in length
     *      48977:  gap of unknown length
     *      48978:  contig of 18 bp in length
     *      48996:  gap of unknown length
     *      48998:  contig of 7309 bp in length
     *      56306:  gap of unknown length
     *      56307:  contig of 18 bp in length
     *      56325:  gap of unknown length
     *      56326:  contig of 7661 bp in length
     *      63986:  gap of unknown length
     *      63987:  contig of 18 bp in length
     *      64005:  gap of unknown length
     *      64006:  contig of 6553 bp in length
     *      70568:  gap of unknown length
     *      70569:  contig of 18 bp in length
     *      70587:  gap of unknown length
     *      70588:  contig of 5814 bp in length
     *      76401:  gap of unknown length
     *      76402:  contig of 18 bp in length
     *      76420:  gap of unknown length
     *      76421:  contig of 10550 bp in length
     *      86970:  gap of unknown length
     *      86971:  contig of 18 bp in length
     *      86989:  gap of unknown length
     *      86990:  contig of 11180 bp in length
     *      98169:  gap of unknown length
     *      98170:  contig of 18 bp in length
     *      98188:  gap of unknown length
     *      98189:  contig of 9964 bp in length
     *     108152:  gap of unknown length
     *     108153:  contig of 18 bp in length
     *     108171:  gap of unknown length
     *     108172:  contig of 11217 bp in length
     *     119388:  gap of unknown length
     *     119389:  contig of 18 bp in length
     *     119407:  gap of unknown length
     *     119408:  contig of 13675 bp in length
     *     133082:  gap of unknown length
     *     133083:  contig of 18 bp in length
     *     133101:  gap of unknown length
     *     133102:  contig of 13241 bp in length
     *     146342:  gap of unknown length
     *     146343:  contig of 18 bp in length
     *     146361:  gap of unknown length
     *     146362:  contig of 17727 bp in length
     *     164088:  gap of unknown length
     *     164089:  contig of 18 bp in length
     *     164106:  gap of unknown length
     *     164107:  contig of 2115 bp in length
     *     166221:  gap of unknown length
     *     166222:  contig of 18 bp in length
     *     166239:  gap of unknown length
     *     166240:  contig of 2341 bp in length
     *     168580:  gap of unknown length
     *     168581:  contig of 18 bp in length
     *     168598:  gap of unknown length
     *     168599:  contig of 2540 bp in length
     *     171138:  gap of unknown length
     *     171139:  contig of 18 bp in length
     *     171156:  gap of unknown length
     *     171157:  contig of 2274 bp in length
     *     173430:  gap of unknown length
     *     173431:  contig of 18 bp in length
     *     173448:  gap of unknown length
     *     173449:  contig of 2030 bp in length
     *     175478:  gap of unknown length
     *     175479:  contig of 18 bp in length
     *     175496:  gap of unknown length
     *     175497:  contig of 2073 bp in length
     *     177569:  gap of unknown length
     *     177570:  contig of 18 bp in length
     *     177587:  gap of unknown length
     *     177588:  contig of 2892 bp in length
     *     180479:  gap of unknown length
     *     180480:  contig of 18 bp in length
     *     180497:  gap of unknown length
     *     180498:  contig of 2413 bp in length
     *     182910:  gap of unknown length
     *     182911:  contig of 18 bp in length
     *     182928:  gap of unknown length
     *     182929:  contig of 2840 bp in length
     *     185768:  gap of unknown length
     *     185769:  contig of 18 bp in length
     *     185786:  gap of unknown length
     *     185787:  contig of 2390 bp in length
     *     188176:  gap of unknown length
     *     188177:  contig of 18 bp in length
     *     188194:  gap of unknown length
     *     188195:  contig of 2490 bp in length
     *     190684:  gap of unknown length
     *     190685:  contig of 18 bp in length
     *     190702:  gap of unknown length
     *     190703:  contig of 3878 bp in length
     *     194580:  gap of unknown length
     *     194581:  contig of 18 bp in length
     *     194598:  gap of unknown length
     *     194599:  contig of 2678 bp in length
     *     197276:  gap of unknown length
     *     197277:  contig of 18 bp in length
     *     197294:  gap of unknown length
     *     197295:  contig of 2701 bp in length
     *     199995:  gap of unknown length
     *     199996:  contig of 18 bp in length
     *     200013:  gap of unknown length
     *     200014:  contig of 2969 bp in length
     *     202982:  gap of unknown length
     *     202983:  contig of 18 bp in length
     *     203000:  gap of unknown length
     *     203001:  contig of 2501 bp in length
     *     205502:  gap of unknown length
     *     205519:  gap of unknown length
     *     205520:  contig of 2180 bp in length
```

```
BASE COUNT     61977 a    41162 c    41291 g    62545 t    724 others
ORIGIN
```

```
Query Match      2.7%;        Score 40.4;   DB 42;   Length 207699;
Best Local Similarity 65.6%;  Pred. No. 19;          Length 207699;
Matches 59;  Conservative 0;  Mismatches 31;  Indels 0;  Gaps 0;
```

RESULT 6

HSDJ130E4/c

```
LOCUS       HSDJ130E4    111944 bp    DNA            HTG       23-NOV-1999
DEFINITION  Homo sapiens chromosome 6 clone RP1-130E4, *** SEQUENCING IN
            PROGRESS ***, in unordered pieces.
ACCESSION   AL078582
VERSION     AL078582.7  GI:6018773
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
```

```
Qy     1455 taaaataaattcaagtattttttaaaaa 1484
            ||||| ||| |||| ||||||| |||| ||
Db   124819 CGTGGGCTTATTTATTTTTATTTTTAAACTATTAATCTTCTGGAATTGCTTTTAGAATTAATTAATG 124760

Qy     1395 cctggatgagtttttttttttgcatcttaataataattcttatatgttgtagaagtatgtatgttt 1454
Db   124759 TAAAAATAATGTAAAATATATTATTTTATAA 124730
```

```
Query Match             2.7%;   Score 40.4;   DB 42;    Length 215566;
Best Local Similarity  65.6%;   Pred. No. 19;
Matches    59;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;
```

```
ORIGIN
BASE COUNT    55527 a    52861 c    53096 g    53425 t      657 others
```

```
FEATURES             Location/Qualifiers
     source          1..215566
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NH0459C22"
```

```
*      138605          gap of unknown length
*      138623 141375:  contig of 2753 bp in length
*      141376          gap of unknown length
*      141393 141394:  contig of 2670 bp in length
*      141394          gap of unknown length
*      144063 144064:  contig of 2188 bp in length
*      144064          gap of unknown length
*      144081 146269:  contig of 3095 bp in length
*      146270          gap of unknown length
*      146287 146288:  contig of 2720 bp in length
*      146288          gap of unknown length
*      149382 149383:  contig of 3473 bp in length
*      149383          gap of unknown length
*      149400 149401:  contig of 2000 bp in length
*      149401          gap of unknown length
*      152120 152121:  contig of 5281 bp in length
*      152121          gap of unknown length
*      152138 152139:  contig of 2716 bp in length
*      152139          gap of unknown length
*      155611 155612:  contig of 3222 bp in length
*      155612          gap of unknown length
*      155629 155630:  contig of 3897 bp in length
*      155630          gap of unknown length
*      155663          gap of unknown length
*      155680 155681:  contig of 3832 bp in length
*      155681          gap of unknown length
*      157629 157630:  contig of 2000 bp in length
*      157630          gap of unknown length
*      157647 157648:  contig of 5281 bp in length
*      157648          gap of unknown length
*      162928 162929:  contig of 3606 bp in length
*      162929          gap of unknown length
*      162946 162947:  contig of 2855 bp in length
*      162947          gap of unknown length
*      165662 165663:  contig of 4219 bp in length
*      165663          gap of unknown length
*      165680 165681:  contig of 3222 bp in length
*      165681          gap of unknown length
*      168902 168903:  contig of 2855 bp in length
*      168903          gap of unknown length
*      168920 168921:  contig of 3606 bp in length
*      168921          gap of unknown length
*      171775 171776:  contig of 4988 bp in length
*      171776          gap of unknown length
*      171793 171794:  contig of 3832 bp in length
*      171794          gap of unknown length
*      175625 175626:  contig of 4970 bp in length
*      175626          gap of unknown length
*      175643 175644:  contig of 3897 bp in length
*      175644          gap of unknown length
*      179540 179541:  contig of 5629 bp in length
*      179541          gap of unknown length
*      179558 179559:  contig of 4219 bp in length
*      179559          gap of unknown length
*      183777 183778:  contig of 6098 bp in length
*      183778          gap of unknown length
*      183795 183796:  contig of 3606 bp in length
*      183796          gap of unknown length
*      187401 187402:  contig of 4988 bp in length
*      187402          gap of unknown length
*      187419 187420:  contig of 6390 bp in length
*      187420          gap of unknown length
*      192407 192408:  contig of 4970 bp in length
*      192408          gap of unknown length
*      192425 192426:  contig of 4988 bp in length
*      192426          gap of unknown length
*      197395 197396:  contig of 6390 bp in length
*      197396          gap of unknown length
*      197413 197414:  contig of 4970 bp in length
*      197414          gap of unknown length
*      203803 203804:  contig of 5629 bp in length
*      203804          gap of unknown length
*      203821 203822:  contig of 3897 bp in length
*      203822          gap of unknown length
*      209450 209451:  contig of 6098 bp in length
*      209451          gap of unknown length
*      209468 209469:  contig of 4219 bp in length
*      209469          gap of unknown length
*      215565 215566:  contig of 6098 bp in length.
*      215566          gap of unknown length.
```

```
Db   124759 TAAAAATAATGTAAAATATATTATTTTATAA 124730
```

RESULT 7

```
REFERENCE   1  (bases 1 to 111944)
  AUTHORS   Parker,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-NOV-1999) Wellcome Trust Genome Campus, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 11, 1999 this sequence version replaced gi:6018475.
            IMPORTANT: This sequence is unfinished and does not necessarily
            represent the correct sequence. Work on the sequence is in progress
            and the release of this data is based on the understanding that the
            sequence may change as work continues. The sequence may be
            contaminated with foreign sequence from E.coli, yeast, vector,
            phage etc. Order of segments is not known; 800 n's separate
            segments. Unfinished: dJ130E4  Contig_ID: 00522  acc=AL078582
            Length: 111944 bp.
            * NOTE: This is a 'working draft' sequence.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
```

```
FEATURES             Location/Qualifiers
     source          1..111944
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP1-130E4"
                     /clone_lib="RPCI-1"
```

```
ORIGIN
BASE COUNT    33294 a    23483 c    22954 g    32213 t
```

```
Qy     1405 tttttttttttgcatcttaataataattcttatatgttgtagaagtatgtatgttt 1464
            ||||| ||| |||| ||||||| ||||| ||  | ||||||||||||||||||||||
Db    40544 TTATATTTATTTAATTAATTAATTAATAAAATAAATAATTAAATTATTTAATTTAAAATAAAT 40485
```

```
Query Match             2.7%;   Score 40.2;   DB 32;    Length 111944;
Best Local Similarity  67.1%;   Pred. No. 19;
Matches    57;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;
```

```
Qy     1465 ttcaagtatttttaaaaaaactt 1489
            |||||||||| ||||||||| |||
Db    40484 ATTAAGTATTTATTTAAAATACATTT 40460
```

RESULT 7

AC011599

```
LOCUS       AC011599    237316 bp    DNA            HTG       21-OCT-1999
DEFINITION  Homo sapiens chromosome 3 seeders clone RPCI11-606L6, ***
            SEQUENCING IN PROGRESS ***, 29 unordered pieces.
ACCESSION   AC011599
VERSION     AC011599.2  GI:6087865
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE   1  (bases 1 to 237316)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodola,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lu,J., Lucier,R.,
            Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M., Morris,S.,
            Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G.,
            Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L., Quiles,M.,
            Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S., Shah,E.,
            Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R., Tabor,P.,
            Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M., Watlington,S.,
```

```
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 237316)
  AUTHORS   Weinstock,G., Weinstock,I.R., Williamson,A., Worley,K., Wren,J.,
            Wrensford,G., Yu,W., Zhou,X., Nelson,D. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Oct 20, 1999 this sequence version replaced gi:6016622.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 29 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1        1094: contig of 1094 bp in length
            *     1095        1961: contig of  867 bp in length
            *     1962        3033: contig of 1072 bp in length
            *     3034        3849: contig of  816 bp in length
            *     3850        4810: contig of  961 bp in length
            *     4811        5605: contig of  795 bp in length
            *     5606        6754: contig of 1149 bp in length
            *     6755        7721: contig of  967 bp in length
            *     7722        8612: contig of  891 bp in length
            *     8613        9661: contig of 1049 bp in length
            *     9662       10580: contig of  919 bp in length
            *    10581       11501: contig of  921 bp in length
            *    11502       12914: contig of 1413 bp in length
            *    12915       13809: contig of  895 bp in length
            *    13810       14425: contig of  816 bp in length
            *    14426       15624: contig of  999 bp in length
            *    15625       16672: contig of 1048 bp in length
            *    16673       17704: contig of 1032 bp in length
            *    17705       19083: contig of 1379 bp in length
            *    19084       20081: contig of  998 bp in length
            *    20082       20958: contig of  877 bp in length
            *    20959       23350: contig of 2392 bp in length
            *    23351       26220: contig of 2870 bp in length
            *    26221       32049: contig of 5829 bp in length
            *    32050       46821: contig of 14772 bp in length
            *    46822       64420: contig of 17599 bp in length
            *    64421       91572: contig of 27152 bp in length
            *    91573      129163: contig of 37591 bp in length
            *   129164      237316: contig of 108153 bp in length.
FEATURES             Location/Qualifiers
     source          1..237316
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="3 seeders"
                     /clone="RPCIll-605L6"
BASE COUNT    54463 a   63949 c   55490 t   75 others
ORIGIN
```

```
Query Match             2.7%;  Score 40.2;  DB 42;  Length 237316;
Best Local Similarity   49.8%;  Pred. No. 21;
Matches  102; Conservative   0; Mismatches  103; Indels   0; Gaps   0;

Qy      746 gaatatgcccaacagagcctgctgcagtgtccgcccggtgactcagtgcagacctgcgggccg 805
            ||| |||||||||||||| ||| | |||||||||| ||| |||||||  || ||||||||||||
Db    58413 GCAGATTCCCACACCCGCGCCTGCCCACTGCCGCCCACGGGAAGTGCCCGGGGAGTGG 58472

Qy      806 gtgcgcttgctccatccatcgtgtctgtgaggaggcctgcagcccaaccggcggactctt 865
            |||  |||||||||| || ||  || || ||||||||  |||   ||||| |||||| |||
Db    58473 AGGCGTGGACAGGCAATCCCGGTGTCTCGAGTCAGCTCAGCCGCCGAGCCGCTTGGACGACAACG 58532

Qy      866 ggttgtggggtgcattctgcagccagtcttcaaggcaagaaacgcaggccagccaggggag 925
            |||| ||||| ||||| |||||| | ||||  || ||||| || | |||||||||| ||||
Db    58533 GCCCGCGGAGGCGAGGCGCAGACAGGGGCGCAGGGCGCCCCCAGGCTCCAGAGCAGCCGCTGAC 58592
```

```
RESULT   8
LOCUS       LEGFRPOL       3417 bp    DNA             PLN       24-AUG-1995
DEFINITION  L.esculentum gene for fruit ripening polygalacturonase.
ACCESSION   X80908
VERSION     X80908.1  GI:963065
KEYWORDS    polygalacturonase.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
            Asteridae; Solananae; Solanales; Solanaceae; Solanum; Potatoe;
            Lycopersicon.
REFERENCE   1 (bases 1 to 3417)
  AUTHORS   Nicholass,F.J., Smith,C.J., Schuch,W., Bird,C.R. and Grierson,D.
  TITLE     High levels of ripening-specific reporter gene expression directed
            by tomato fruit polygalacturonase gene-flanking regions
  JOURNAL   Plant Mol. Biol. 28 (3), 423-435 (1995)
  MEDLINE   95359401
REFERENCE   2 (bases 1 to 3417)
  AUTHORS   Nicholass,F.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-1994) F.J. Nicholass, University of Nottingham,
            AFRC Research Group in Plant Gene Reg., Dept of Physiology & Envir.
            Science, Sutton Bonigton Campus, Loughborough Leics. LE12 5RD, UK
FEATURES             Location/Qualifiers
     source          1..3417
                     /organism="Lycopersicon esculentum"
                     /cultivar="Ailsa Craig"
                     /db_xref="taxon:4081"
                     /chromosome="10"
                     /clone="GTOM23.73"
     mRNA            1..3417
                     /product="fruit ripening specific polygalacturonase"
BASE COUNT     1307 a    438 c    424 g   1248 t
ORIGIN
```

```
Query Match             2.7%;  Score 40;  DB 7;  Length 3417;
Best Local Similarity   65.9%;  Pred. No. 14;
Matches   58; Conservative   0; Mismatches  30; Indels   0; Gaps   0;

Qy     1400 atgagttttttttttgcatcttaatattcttatatgttgagagtgatgtttaaaa 1459
            |||| ||| || ||||| |||  | || ||| ||  || ||||| |||| ||||||||
Db     3173 AAGAATTATTTTTGAGCAACAAATTATATATTTAGTCACGCTTTTAAGTGGTATTTTTAAAA 3232

Qy     1460 taaattttcaagtattttttaaaaaact 1487
            |||| || || |||||| |||||| ||||
Db     3233 TAAAAATTAAGGTATTATTTGAAAAAAATT 3260
```

```
RESULT   9
LOCUS       AC002449       173767 bp    DNA             PRI       20-AUG-1997
DEFINITION  Human PAC clone DJ404K21 from Xq23, complete sequence.
ACCESSION   AC002449
VERSION     AC002449.1  GI:2337886
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 173767)
  AUTHORS   Hinds,K, Tin-Wollam,A, Sutterer,C and Fronick,B.
  TITLE     The sequence of H. sapiens PAC clone DJ404K21
  JOURNAL   Unpublished (1997)
REFERENCE   2 (bases 1 to 173767)
```

```
AUTHORS   Waterson,R.
TITLE     Direct Submission
JOURNAL   Submitted (20-AUG-1997) Department of Genetics, Washington
          University 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
SUBMITTED BY:
          Genome Sequencing Center
          Department of Genetics
          Washington University
          St. Louis MO 63108, USA
          http://genome.wustl.edu/gsc
          mailto:saplens@watson.wustl.edu

COMMENT   NOTICE: This sequence may not represent the entire insert of this
          clone.  It may be shorter because we only sequence overlapping
          clone sections once, or longer because we provide a small overlap
          between neighboring clone submissions.

          This sequence was finished as follows unless otherwise noted:
          all regions were double stranded or sequenced with an alternate
          chemistry; an attempt was made to resolve all sequencing problems,
          such as compressions and repeats; all regions were covered by
          sequence from more than one subclone; and the assembly was
          confirmed by restriction digest.

          MAPPING INFORMATION:
          This sequence was generated from part of bacterial clone contigs of
          human chromosome X, constructed by David Bentley's chromosome X
          mapping group at the Sanger Centre, Wellcome Trust Genome Campus,
          Hinxton, UK.  Further information can be found at
          http://www.sanger.ac.uk/HGP/ChrX

          SOURCE INFORMATION:
          This clone was derived from human PAC library RPCI-3 prepared by
          Pieter de Jong and coworkers at Roswell Park Cancer Institute,
          using the method described by Ioannou et al., Nature Genetics
          6:84-9 (1994).  The library is from one male donor.  For further
          details, see http://bacpac.med.buffalo.edu/  The clone is available
          from Genome Systems, Inc. (http://www.genomesystems.com).
          VECTOR: pCYPAC2

          NEIGHBORING SEQUENCE INFORMATION:
          The clone sequenced to the left is H_DJ290B04.  The actual start of
          this clone is at base position 1 of DJ404K21; actual end is at
          173767 of DJ404K21. This clone is part of an unanchored island,
          orientation is unknown.

          This clone contains STS's swXD757 (NID:g405440) and swXD512
          (NID:g995449).

FEATURES             Location/Qualifiers
     source          1..173767
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /clone="DJ404K21"
                     /clone_lib="RPCI-3"
                     /map="Xq23"
     repeat_region   3958..3983
                     /rpt_family="L1"
     repeat_region   7187..7448
                     /rpt_family="MER"
     repeat_region   7450..7832
                     /rpt_family="ALU"
     repeat_region   complement(7833..7891)
                     /rpt_family="ALU"
     repeat_region   7893..8184
                     /rpt_family="THR"
     repeat_region   complement(8198..9670)
                     /rpt_family="ALU"
     repeat_region   9671..10045
                     /rpt_family="THR"
     repeat_region   10047..10226
                     /rpt_family="ALU"
     repeat_region   10461..11065
                     /rpt_family="MER"
     repeat_region   complement(11071..11286)
                     /rpt_family="L1"
     repeat_region   11291..11433
                     /rpt_family="MER"
     repeat_region   complement(11434..11593)
                     /rpt_family="L1"
     repeat_region   complement(11616..11641)
                     /rpt_family="MER"
     repeat_region   complement(11472..14820)
                     /rpt_family="MER"
     repeat_region   complement(14833..15202)
                     /rpt_family="L1"
     repeat_region   15777..15808
                     /rpt_family="L1"
     repeat_region   16030..16436
                     /rpt_family="L1"
     repeat_region   16625..16719
                     /rpt_family="L1"
     repeat_region   16860..18448
                     /rpt_family="L1"
     repeat_region   18450..18738
                     /rpt_family="L1"
     repeat_region   18755..22756
                     /rpt_family="ALU"
     repeat_region   complement(24561..24819)
                     /rpt_family="L1"
     repeat_region   25322..25360
                     /rpt_family="ALU"
     repeat_region   25871..25936
                     /rpt_family="L1"
     repeat_region   27205..27559
                     /rpt_family="L1"
     repeat_region   27921..28394
                     /rpt_family="ALU"
     repeat_region   29212..29235
                     /rpt_family="L1"
     repeat_region   complement(30468..30557)
                     /rpt_family="L1"
     repeat_region   30844..31126
                     /rpt_family="L1"
     misc_feature    31050..31156
                     /note="match to EST R00201 (NID:g749937) ye71f11.r1"
     misc_feature    31530..32115
                     /note="match to EST R00201 (NID:g749937) ye71f11.r1"
     repeat_region   complement(34006..34295)
                     /rpt_family="L1"
     repeat_region   35839..35865
                     /rpt_family="ALU"
     repeat_region   36463..37291
                     /rpt_family="L1"
     repeat_region   39129..39159
                     /rpt_family="L1"
     repeat_region   complement(39829..40254)
                     /rpt_family="L1"
     repeat_region   40512..41114
                     /rpt_family="ALU"
     repeat_region   41227..41458
                     /rpt_family="L1"
     repeat_region   complement(41461..41750)
                     /rpt_family="L1"
     repeat_region   42582..42630
                     /rpt_family="ALU"
     repeat_region   43534..43661
                     /rpt_family="L1"
     repeat_region   44290..44440
                     /rpt_family="L1"
     repeat_region   44494..44678
                     /rpt_family="L1"
     repeat_region   complement(45130..45422)
                     /rpt_family="ALU"
     repeat_region   complement(45972..46009)
                     /rpt_family="THR"
```

```
repeat_region    47644..47668
                 /rpt_family="L1"
repeat_region    48498..48788
                 /rpt_family="ALU"
repeat_region    complement(49482..49770)
                 /rpt_family="ALU"
repeat_region    complement(51436..51509)
                 /rpt_family="L1"
repeat_region    complement(51533..51623)
                 /rpt_family="L1"
repeat_region    complement(52389..52862)
                 /rpt_family="L1"
repeat_region    complement(52874..54009)
                 /rpt_family="L1"
repeat_region    complement(54244..54351)
                 /rpt_family="L1"
repeat_region    complement(54308..54351)
                 /rpt_family="L1"
repeat_region    54388..54524
                 /rpt_family="ALU"
repeat_region    54525..57141
                 /rpt_family="ALU"
repeat_region    57168..58829
                 /rpt_family="L1"
repeat_region    58231..58520
                 /rpt_family="ALU"
repeat_region    58522..61300
                 /rpt_family="ALU"
repeat_region    61313..61482
                 /rpt_family="L1"
repeat_region    complement(61490..61545)
                 /rpt_family="L1"
repeat_region    complement(61550..61830)
                 /rpt_family="ALU"
repeat_region    complement(61833..61952)
                 /rpt_family="ALU"
repeat_region    63130..63422
                 /rpt_family="L1"
repeat_region    complement(63435..63457)
                 /rpt_family="MER"

Query Match    Best Local Similarity  55.9%;
Matches  76;  Conservative  2.78;  Pred. No. 23;  DB 10;  Length 173767;
Mismatches  60;  Indels  0;  Gaps  0;

Qy      43 ataaatacacattgataagaaagatggcttaaaagtgctactaagaacaagaagaaaacgtt 102
           ||||||||||   | | |  |||||||  |||    ||| || |||  ||||| | |||||
Db   39644 AAAAAATACATTTAAAAAGAAGAGGGGATGACAAGAAATACACTAATAATAAAAAGACATT 39703

Qy     103 tttcactctttagtattactaggctattgtcatgtaaagtgacttgtgaatcaggaga 162
           | | | | |||  |  ||  |||| || |||  | |||| | ||||| | |||||
Db   39704 AATTTGTCAAATTCCAAGACAGCAGGGTCGAGTAGACAAGCAAGTAAGGATATAAGAGAGACTCAAATGA 39763

Qy     163 ctgtgagacagcaagaa 178
           | |||    |  ||||
Db   39764 CGTAATTAATGAAGCA 39779
```

RESULT 10
```
LOCUS        D88490        2670 bp    mRNA        PLN    26-JUN-1999
DEFINITION   Candida albicans mRNA for CaKRE6, complete cds.
ACCESSION    D88490
VERSION      D88490.1  GI:1944423
KEYWORDS     CaKRE6.
SOURCE       Candida albicans cDNA to mRNA.
ORGANISM     Candida albicans
             Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
             Candidaceae; Candida.
REFERENCE    1  (sites)
AUTHORS      Mio,T., Yamada-Okabe,T., Yabe,T., Nakajima,T., Arisawa,M. and
             Yamada-Okabe,H.
TITLE        Isolation of the Candida albicans homologs of Saccharomyces
JOURNAL      cerevisiae KRE6 and SKN1: expression and physiological function
             J. Bacteriol. 179 (7), 2363-2372 (1997)
MEDLINE      97234650
REFERENCE    2  (bases 1 to 2652)
AUTHORS      Mio,T.; Yamada-Okabe,T., Yabe,T., Nakajima,T., Arisawa,M. and
             Yamada-Okabe,H.
TITLE        Isolation from Candida albicans of the Saccharomyces cerevisiae
             KRE6 and SKN1homologues: expression and physiological function
JOURNAL      Unpublished (1997)
REFERENCE    3  (bases 1 to 2670)
AUTHORS      Okabe,T.
TITLE        Direct Submission
JOURNAL      Submitted (21-OCT-1996) to the DDBJ/EMBL/GenBank databases. Toshiko
             Okabe, Yokohama City University, School of Medicine, Department of
             Hygiene; 3-9 Fukuura, Kanazawa-ku, Yokohama, Kanagawa 236-0004,
             Japan (E-mail:toshiko@med.yokohama-cu.ac.jp, Tel:81-45-787-2606,
             Fax:81-45-786-5219).
FEATURES             Location/Qualifiers
     source          1..2670
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
     CDS             51..2273
                     /codon_start=1
                     /transl_table=12
                     /product="CaKRE6"
                     /protein_id="BAA19593.1"
                     /db_xref="GI:1944424"
                     /translation="MASQRDLTSNNPHFHISSSRNNTNDDFNSDHNPRNPEFGDPEEE
                     EELDSASFSSSQNQPHHFPATSNQTSSTNNLHSTNSIRNEAEYGKPFQAYSGIYSN
                     GGSSSYLNDEGVISDESRLLSSGNNNNNNNNNRNNHQHPHLGSSVSVGGGGGDNRLNPTSPAEFD
                     RYPSMAGSRVVSSTSLVSFNNPSNMMRNHIQHPVQDAEYEKNRPMHDLKNMDKRSL
                     PFPNDFSPFGGYPASSFPLSIDEKEPDDYLNNPDPVQDAEYEKNRPMHDLKNMDKRSL
                     DTPESAYKKAKDGSEWVLVFSDEFDAEGRFFYEGDDGFTAPDIHYDATKDLEWDP
                     GGLGFIILPIAALAVFILPALIYSGANPYHPRSYEVLFLTKYSYPMLSAIRMNLVDP
                     DAVTTANGTLNLRMDAYKNNHLFYRSGMVQSWNQLCYTQGHLEISARLPNYGNVTGLW
                     PGLWSMGNLGRPGYLGSTDGVWPYSYDSCDAGITPNQSSPDGISYLPGQRLNKCTCPG
                     ELHPNRGVGRGAPEIDVIEGEVMTDSSGKKENCGVASQSLQLAPMDIWYIPDYNNVEI
                     YNFSVSTMNYTGGPPQQALSATTMLNVTWTYEFGDNAHNFQTYGYEYLNDPETGYLRW
                     FVGDDPTLVYSQALHPDGNIGWRPLSKEPMSLIILNLGISNNWAYIDWPSISFPVTFR
                     IDYVRVYQPPDQINVGCDPIDFPTYDYIQQHLNLYENANITSFEDGGYKFPKNSLIGC
     polyA_site      2652
                     /note="18 A nucleotides"
BASE COUNT      874 a    404 c    457 g    935 t
ORIGIN
```

RESULT 11
```
LOCUS        PFASSA        4673 bp    DNA        INV    26-APR-1993
DEFINITION   Plasmodium yoelii sporozoite surface protein 2 gene, complete cds.
ACCESSION    M84732
VERSION      M84732.1  GI:160692
KEYWORDS     sporozoite surface protein.
SOURCE       Plasmodium yoelii erythrocytic stage blood DNA.
ORGANISM     Plasmodium yoelii
             Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE    1  (bases 1 to 4673)
AUTHORS      Rogers,W.O., Rogers,M.D., Hedstrom,R. and Hoffman,S.L.
TITLE        Characterization of the gene encoding Sporozoite Surface Protein 2,

Query Match    Best Local Similarity  67.5%;
Matches  56;  Conservative  2.78;  Pred. No. 15;  DB 7;  Length 2670;
Mismatches  27;  Indels  0;  Gaps  0;

Qy    1403 agtttttttttggcatcctttaaataatttcttaatatgttgtagagtatgttttaaaataa 1462
           |||| ||||||||    |||||    || || |||||  |||| ||| | || |||||||| ||
Db    2579 AGTCTTTTTTTATTATTATTTAAATATTATATATTATTTTTTTTAAATAAA 2638

Qy    1463 atttcaagtatttttttaaaaa 1485
           || |||| |||| || |||||
Db    2639 GCTTTTATTTTATTCAAAAAAAA 2661
```

JOURNAL   a protective Plasmodium yoelii sporozoite antigen
MEDLINE   Mol. Biochem. Parasitol. 53, 45-52 (1992)
             92365736
FEATURES   Location/Qualifiers
  source   1..4673
             /organism="Plasmodium yoelii"
             /db_xref="taxon:5861"
             /cell_type="erythrocytic stage"
             /dev_stage="erythrocytic stage"
             /tissue_type="blood"
  gene   718..3198
             /gene="sporozoite surface protein"
  CDS   718..3198
             /gene="sporozoite surface protein"
             /codon_start=1
             /product="sporozoite surface protein"
             /protein_id="AAA29768.1"
             /db_xref="GI:160693"
             /translation="MKLLGNSKYIFVVLLLCISVFLNGQETLDEIKYSEEVCTEQIDI
             HILLDGSGSIGYSNWKAHVIPMLNTLVDNLNISNDEINVSLLFSTNSRELIKLKGYG
             STSKSDLRFIALHLQNNYSPNGNTNLTSALLVVDTLINERMYRPDAIQLAIILTDGIP
             NDLPRSTAVVHQLKRKHVNVAIIGVGAGVNEYNRILVGCDRYAPCPYYSSGSWNEAQ
             NMIKPFLTKVCQEVERIAHCGKWEEWSECSTTCDEGRIRRQILHPGCVSEMTTPCK
             VRDCPQIPIPIVPINKIPEKPSNPEEEVYNPDPNDPNPNPNPNPNPNPNPNPNPNPNPN
             PNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNPNDPSPNPNPNPHPK
             RNPKRNPDNKPKPNKPNPNKPNPNEPSPNPNKPNPNKPNPNPNPNPNPNPNPNPNPNPN
             EPSNPRKPSNPRKPNPNEPSNPNPAPSPNPNPNPNPNPNPNPNPNPNPNPNPNEPSNPNP
             NPEENPEEPSNPNKPEEPSNPNEPLNPNEPNPNPNPNPNPNPNPNPNPSNPNEPSNPNEP
             NPKPSNPNKPSNPKEPSNPNEPNPSNPNKPNPNPNPNPNPNPNEPNPNEPSNPNEPSNP
             EDNENPLLIIQDDPIEPRNDSNVIPLIPIPQKGNNIPSNLPENPSDSEVEYPRPNDNG
             ENSNNTMKSKRNIPNEFIPSPGDNPYKGHEERIPKPHRSNDIYIDNNVNKNKDEPE
             IPNEYEEDKNKNQSKSNNGYKIAGGIIGGLIALLGCAGVGYNPIAGSSAAGLAGAEPA
             PEDVIPDDKDIVENEQFKLPEDNDWN"

BASE COUNT   1836 a   807 c   530 g   1500 t
ORIGIN

RESULT 12
ACC006803/c
LOCUS   AC006803   299820 bp   DNA   HTG   25-FEB-1999
DEFINITION   Caenorhabditis elegans clone Y53G8A, *** SEQUENCING IN PROGRESS
             ***, 3 unordered pieces.
ACCESSION   AC006803
VERSION   AC006803.2   GI:4309783
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   Caenorhabditis elegans.
ORGANISM   Caenorhabditis elegans
             Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
             Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1 (bases 1 to 299820)
AUTHORS   Waterston,R.H.
TITLE   The sequence of Caenorhabditis elegans clone
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 299820)
AUTHORS   Waterston,R.H.
TITLE   Direct Submission
JOURNAL   Submitted (23-FEB-1999) Genome Sequencing Center, Washington
             University School of Medicine, 4444 Forest Park Parkway, St. Louis,
             MO 63108, USA
COMMENT   On Mar 1, 1999 this sequence version replaced gi:4263135.

Query Match   2.7%;   Score 39.6;   DB 34;   Length 4673;
Best Local Similarity   60.0%;   Pred. No. 19;
Matches   66;   Conservative   0;   Mismatches   44;   Indels   0;   Gaps   0;

Qy   1379 ataagatttggggaacctggatgagtcgagtctttttttttcatcctttaataatttcttatat 1438
           | |||| ||||| ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   4033 ATAATATTATTGTGTAAAAATTAAGATGATATTATTTTTAGACATATTTGACAAATTGTCAAAT 4092

Qy   1439 gttgtagagtatgttttaaaataaattccaagtatttttttaaaaaactt 1488
           | |||| ||| ||||||||||||||| |||||||||||||||||||||||
Db   4093 TCGTTGTTATTCTGTTTTTGTAAAAATATGAGCATTTTTTTATAATTATT 4142

RESULT 13
ACC006871/c
LOCUS   AC006871   299820 bp   DNA   HTG   26-FEB-1999
DEFINITION   Caenorhabditis elegans clone Y53G8X, *** SEQUENCING IN PROGRESS
             ***, 3 unordered pieces.
ACCESSION   AC006871
VERSION   AC006871.1   GI:4263477
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   Caenorhabditis elegans.
ORGANISM   Caenorhabditis elegans
             Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
             Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1 (bases 1 to 299820)
AUTHORS   Waterston,R.H.
TITLE   The sequence of Caenorhabditis elegans clone
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 299820)
AUTHORS   Waterston,R.H.
TITLE   Direct Submission
JOURNAL   Submitted (24-FEB-1999) Genome Sequencing Center, Washington
             University School of Medicine, 4444 Forest Park Parkway, St. Louis,
             MO 63108, USA
COMMENT   * NOTE: This is a 'working draft' sequence. It currently
             * consists of 3 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *   1   76548: contig of 76548 bp in length
             *   76549   76557: gap of unknown length
             *   76558   162292: contig of 85735 bp in length
             *   162293   162300: gap of unknown length
             *   162301   299820: contig of 137520 bp in length.
FEATURES   Location/Qualifiers
  source   1..299820
             /organism="Caenorhabditis elegans"

BASE COUNT   98291 a   51999 c   51839 g   96253 t   1438 others
ORIGIN

Query Match   2.7%;   Score 39.6;   DB 33;   Length 299820;
Best Local Similarity   60.0%;   Pred. No. 31;
Matches   66;   Conservative   0;   Mismatches   44;   Indels   0;   Gaps   0;

Qy   1380 taagatttggggaacctggatgagtcgagtctttttttttcatcctttaataatttcttatatg 1439
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 159048 TGATAATTATAGTAAAACTCGACGAGTTTTGAAAAGGTTTTCAGGCAAATTTTTTTCAAAT 158989

Qy   1440 ttgtagagtatgttttaaaataaatttcaagtatttttttaaaaaactt 1489
           ||||| |||| ||||||||||||| ||||||||||||||| ||||||||||
Db 158988 ATTCGGAAAAATTTTGTTTCAAAAAATTTTAAAAATAATTTTTTAAAAACAAATTT 158939

Qy   1379 ataagatttggggaacctggatgagtcgagtctttttttttcatcctttaataatttcttatat 1438
Db   4033 ATAATATTATTGTGTAAAAATTAAGATGATATTATTTTTAGACATATTTGACAAATTGTCAAAT 4092

```
BASE COUNT    98291 a   51999 c   51839 g   96253 t   1438 others
ORIGIN

Qy      1380  taagaattgggggaacctggatggagtcttttttttttttttgcatctttaataattcttatatg 1439
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    159048  TGATAATTAGAGTAAAACTCGACGACGAGTTTTGAAAGGTTTTCAGGCAAATTTTTTTCAAAAT 158989

Qy      1440  ttgtagagtatgttttaaaaataaattcaagtattttttaaaaaaacctt 1489
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    158988  ATTCGGAAAATTTGTTTCAAAAAATTTAAAAATTAATTTTTTTAAAACAATTT 158939
```

```
RESULT 14
DDU73686
LOCUS       DDU73686              1840 bp    DNA             INV       08-DEC-1997
DEFINITION  Dictyostelium discoideum cytosolic glycoprotein FP21 (fpa2) gene,
            complete cds.
ACCESSION   U73686
VERSION     U73686.1  GI:1658023
KEYWORDS    .
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Dictyostelida; Dictyostelium.
REFERENCE   1  (bases 1 to 1840)
  AUTHORS   West,C.M., Kozarov,E. and Teng-unmuay,P.
  TITLE     The cytosolic glycoprotein FP21 of Dictyostelium discoideum is
            encoded by two genes resulting in a polymorphism at a single amino
            acid position
  JOURNAL   Gene 200 (1-2), 1-10 (1997)
  MEDLINE   98038971
REFERENCE   2  (bases 1 to 1840)
  AUTHORS   West,M.C., Kozarov,E. and Teng-unmuay,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-1996) Anatomy & Cell Biology, University of
            Florida, 1600 SW Archer Rd., Gainesville, FL 32610-0235, USA
FEATURES             Location/Qualifiers
     source          1..1840
                     /organism="Dictyostelium discoideum"
                     /strain="Ax3"
                     /db_xref="taxon:44689"
     gene            1061..1703
                     /gene="fpa2"
     CDS             join(1061..1157,1312..1703)
                     /gene="fpa2"
                     /note="SKR1-like; similar to the product of the fpa1 gene"
                     /codon_start=1
                     /product="cytosolic glycoprotein FP21"
                     /protein_id="AAB88390.1"
                     /db_xref="GI:1658024"
                     /translation="MSLVKLESSDEKVFEIEKEIACMSVYIKNMEDIGESDAPIPLP
                     NVISTILEKVLDYCRHHQHPSPQGDDKKDEKRLDDIPPYDRDCKVDQPTLFELILA
                     ANYLDIKPLLDVTCKTVAANMIRGKTPEEIRKIFNIKNDFTPEEEEQIRKENEWCEDKG
                     GN"
     intron          1158..1311
                     /gene="fpa2"
BASE COUNT      742 a     166 c     170 g     762 t
ORIGIN

Qy      1393  aacctggatgagtttttttgcatctttaataataatgttggtagagtatgt 1452
              ||| || |||||||||||||||||||| |||| |||||||||||| |||||||
Db       870  AACTTTTTTTTTTTTTTTTTTTTTTTTTTCTATTTTTTTTTAATTTTTTTTTTTTTTTTTT 929
```

```
Qy      1453  tttaaaataaattcaagtattttttaaaaaaacctt 1489
              |||||||| ||| || |||| || | ||| ||| |||||| |||||| |||
Db       930  TTTAAAAATAAAATTAATTAATTTTTTTTTTTTAAAAAAAGTTTT 966
```

```
RESULT 15
AC004994/c
LOCUS       AC004994            80331 bp    DNA             PRI       21-NOV-1998
DEFINITION  Homo sapiens PAC clone DJ1194E15 from 7p15.3-p21, complete
            sequence.
ACCESSION   AC004994
VERSION     AC004994.1  GI:3900849
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 80331)
  AUTHORS   Drone,K., Ryan,E. and Williamson,A.
  TITLE     The sequence of Homo sapiens PAC clone DJ1194E15
  JOURNAL   Unpublished (1998)
REFERENCE   2  (bases 1 to 80331)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 80331)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Nov 21, 1998 this sequence version replaced gi:3377480.
            SUBMITTED BY: WUGSC
            Genome Sequencing Center
            Department of Genetics
            Washington University
            St. Louis MO 63108; USA
            http://genome.wustl.edu/gsc
            mailto:saplens@watson.wustl.edu

            NOTICE:   This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded or sequenced with an alternate
            chemistry; an attempt was made to resolve all sequencing problems,
            such as compressions and repeats; all regions were covered by
            sequence from more than one subclone; and the assembly was
            confirmed by restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center.  For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            This clone was derived from human PAC library RPCI-5, prepared by
            Pieter de Jong and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu) using the method described by
            Ioannou et al., Nature Genetics 6:84-9 (1994).  The library is from
            one male donor.
            The clone may be obtained either from Genome Systems, Inc.
            (http://www.genomesystems.com) or Research Genetics, Inc.
            (http://www.resgen.com); or from Pieter de Jong.
```

```
BASE COUNT
ORIGIN
     /db_xref="taxon:6239"
     /clone="Y53G8X"

Qy         a   51999 c

Query Match    2.7% ;
Best Local Similarity    60.0% ;
Matches   66;   Conservative    0;   Pred. No. 31;

              Score 39.6;   DB 33;   Length 299820;
              Mismatches   44;   Indels    0;   Gaps    0;

Query Match    2.68% ;
Best Local Similarity    62.9% ;
Matches   61;   Conservative    0;   Mismatches   36;   Indels    0;   Gaps    0;

              Score 39.4;   DB 34;   Length 1840;
              Pred. No. 18;
```

```
VECTOR: pCYPAC2
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is GS465N13, 200 bp overlap; the
clone sequenced to the right is RG264L19, 200 bp overlap.  Actual
start of this clone is at base position 76644 of GS465N13; actual
end is at 4319 of RG264L19.

FEATURES             Location/Qualifiers
     source          1..80331
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7p15.3-p21"
                     /clone_lib="RPCI-5"
                     /clone="DJ1194E15"
     gene            complement(<13..>79764)
                     /gene="WUGSC:H_DJ1194E15.1"
     CDS             complement(join(<13..97,18301..18434,79647..>79764))
                     /gene="WUGSC:H_DJ1194E15.1"
                     /note="similar to KIAA0600; similar to d1026456
                     (PID:g3043724); H_DJ1194E15.1"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAC78618.1"
                     /db_xref="GI:3900850"
                     /translation="MIVKPVAKEPDPDMVLVSAGFDALEGHTPPLGGYKYTAKCFGHL
                     TKQLMTLADGHDLTAICDASEACVNALLGNELEPLAEDILHQSPNMNAV
                     ISLOKIIEIQ"
     repeat_region   2455..2512
                     /rpt_family="AT_rich"
     repeat_region   3372..3565
                     /rpt_family="L1"
     repeat_region   4099..4315
                     /rpt_family="L2"
     repeat_region   4316..4365
                     /rpt_family="L1"
     repeat_region   4483..5333
                     /rpt_family="(TA)n"
     misc_feature    4958..5023
                     /gene="WUGSC:H_DJ1194E15.1"
                     /note="nucleolar phosphoprotein B23 pseudogene"
     repeat_region   5656..5753
                     /rpt_family="(CAT)n"
     repeat_region   5853..7796
                     /rpt_family="L1"
     repeat_region   7927..8176
                     /rpt_family="L1"
     repeat_region   8177..8486
                     /rpt_family="L1"
     repeat_region   8487..9690
                     /rpt_family="Alu"
     repeat_region   11262..11310
                     /rpt_family="L1"
     repeat_region   11381..11577
                     /rpt_family="AT_rich"
     repeat_region   12987..13102
                     /rpt_family="MER1_type"
     repeat_region   15624..15929
                     /rpt_family="purine-rich"
     repeat_region   15989..16069
                     /rpt_family="Alu"
     repeat_region   16906..17212
                     /rpt_family="Alu"
     repeat_region   17811..17894
                     /rpt_family="Alu"
     repeat_region   18501..18681
                     /rpt_family="MIR"
     repeat_region   18854..18950
                     /rpt_family="MIR"
     repeat_region   19357..19427
                     /rpt_family="AT_rich"
     repeat_region   1973..19619
                     /rpt_family="L1"
     repeat_region   19631..19916
                     /rpt_family="L1"
```

```
     repeat_region   20235..20524
                     /rpt_family="L1"
     repeat_region   20590..20621
                     /rpt_family="L1"
     repeat_region   20683..20891
                     /rpt_family="AT_rich"
     repeat_region   20896..21144
                     /rpt_family="MaLR"
     repeat_region   21364..21482
                     /rpt_family="MaLR"
     repeat_region   22050..22083
                     /rpt_family="purine-rich"
     repeat_region   22088..22247
                     /rpt_family="AT_rich"
     repeat_region   23613..23634
                     /rpt_family="AT_rich"
     repeat_region   23754..23794
                     /rpt_family="(CA)n"
     repeat_region   24294..24587
                     /rpt_family="Alu"
     repeat_region   24901..24939
                     /rpt_family="AT_rich"
     repeat_region   25243..25281
                     /rpt_family="AT_rich"
     repeat_region   25327..25336
                     /rpt_family="AT_rich"
     repeat_region   25395..25457
                     /rpt_family="AT_rich"
     repeat_region   25576..25622
                     /rpt_family="L2"
     repeat_region   27713..27849
                     /rpt_family="(TGGA)n"
     repeat_region   27850..27880
                     /rpt_family="MIR"
     repeat_region   27889..28176
                     /rpt_family="L2"
     repeat_region   28711..28846
                     /rpt_family="Alu"
     repeat_region   28829..28909
                     /rpt_family="L2"
     repeat_region   31608..31713
                     /rpt_family="MIR"
     repeat_region   31897..31947
                     /rpt_family="MER1_type"
     repeat_region   32013..33329
                     /rpt_family="AT_rich"
     repeat_region   33329..34415
                     /rpt_family="L1"
     repeat_region   35816..35997
                     /rpt_family="L1"
     repeat_region   36396..36522
                     /rpt_family="L2"
     repeat_region   37800..37879
                     /rpt_family="MIR"
     repeat_region   38441..38463
                     /rpt_family="MIR"
     repeat_region   39574..39889
                     /rpt_family="(CAA)n"
     repeat_region   39911..40294
                     /rpt_family="MER1_type"
     repeat_region   41534..41740
                     /rpt_family="MaLR"
```

```
Query Match                   2.6%;  Score 39.4;  DB 40;  Length 80331;
Best Local Similarity   56.6%;  Pred. No. 29;
Matches    73;  Conservative   0;  Mismatches  56;  Indels   0;  Gaps   0;

Qy    1357 agatcaagaactctgtcccagcataagatttcggggaacctggatgagtcttttttttg 1416
           ||||||| ||| ||||||||||| |||| ||||||||||||||||||| |||| |||
Db   77646 AAATCAAAAGTTCTTATCCTGAAACAGTAGTTAGGTGAAATTGCATGAGCCATTTTCTGTA 77587

Qy    1417 catcttaataatttcttatatgttgtagagtatgtttaaaataaatttcaagtatt 1476
           ||  || | |  |||  | |||  ||| ||| | |||  ||||||   |||||
```

```
Db  77586 TTTTTTCCATTAAAAAGAAAAATGTTTCATAAAAATGGATGAGAATAGATTTTTTTTTTTTT 77527

Qy   1477 tttaaaaaa 1485
          ||||| | |
Db  77526 TTTAAGACA 77518
```

Search completed: April 23, 2000, 02:20:29
Job time: 3094 sec

MPSRCH (TM)

```
***************************************************************
Release 3.1A   John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd
***************************************************************
```

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 18 14:00:52 2000;   MasPar time 93.97 Seconds
           312.120 Million cell updates/sec

Tabular output not generated.

Post-processing: Minimum Match 0%
                    Listing first 45 summaries

Searched: 225878 seqs, 69334122 residues

Scoring table: PAM 150
                  Gap 11

Sequence: 1 MALKVLLEQEKTFLLVLL.........AVIHPATQTSLQVRQRLGSL 423
Perfect Score: 3111
Description: (1-423) from US09490187.pep
Title: >US-09-490-187-2

Database: sptrembl12
    1:sp_archea 2:sp_bacteria 3:sp_fungi 4:sp_human
    5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
    9:sp_plant 10:sp_plant 11:sp_rodent 12:sp_unclassified
    13:sp_vertebrate 14:sp_virus

Statistics: Mean 46.756;   Variance 76.185;   scale 0.614

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match %	Length	DB	ID	Description	Pred. No.
1	205	6.6	186	14	O72735	A53R PROTEIN.	8.55e-21
2	201	6.5	186	14	Q9WJB4	TUMOUR NECROSIS FACTOR	5.80e-20
3	187	6.0	186	14	Q9YP87	TUMOUR NECROSIS FACTOR	4.32e-17
4	161	5.2	283	6	Q9XSZ8	SOLUBLE TNF RECEPTOR C	6.13e-12
5	152	4.9	348	14	O57108	HVEAS.	6.13e-10
6	152	4.9	348	14	O57110	TUMOR NECROSIS FACTOR	3.19e-10
7	152	4.9	349	14	O57277	TUMOR NECROSIS FACTOR	3.19e-10
8	152	4.9	349	14	O57103	TUMOR NECROSIS FACTOR	3.19e-10
9	152	4.9	349	14	O57102	TUMOR NECROSIS FACTOR	3.19e-10
10	152	4.9	349	14	O57101	TUMOR NECROSIS FACTOR	3.19e-10
11	152	4.9	349	14	O57100	TUMOR NECROSIS FACTOR	3.19e-10
12	152	4.9	349	14	O57099	TUMOR NECROSIS FACTOR	3.19e-10
13	150	4.8	253	4	O00276	LYMPHOCYTE ASSOCIATED	7.58e-10
14	150	4.8	277	4	O14866	SOLUBLE DEATH RECEPTOR	7.58e-10
15	150	4.8	350	14	O57116	TUMOR NECROSIS FACTOR	7.58e-10
16	149	4.8	350	14	O57123	TUMOR NECROSIS FACTOR	1.17e-09
17	148	4.8	355	14	Q85308	SECRETED RECEPTOR BIND	1.79e-09
18	150	4.8	372	4	O00279	LYMPHOCYTE ASSOCIATED	7.58e-10
19	150	4.8	380	4	O00280	LYMPHOCYTE ASSOCIATED	7.58e-10
20	150	4.8	418	4	O00275	LYMPHOCYTE ASSOCIATED	7.58e-10
21	150	4.9	426	4	O14865	DEATH RECEPTOR 3 BETA.	7.58e-10
22	143	4.6	326	14	O57120	TUMOR NECROSIS FACTOR	1.51e-08
23	144	4.6	351	14	O73559	H4R PROTEIN.	9.90e-09
24	144	4.6	326	14	O57122	TUMOR NECROSIS FACTOR	3.52e-08
25	140	4.5	347	14	O57115	TUMOR NECROSIS FACTOR	5.35e-08
26	141	4.5	349	14	O57305	TUMOR NECROSIS FACTOR	3.52e-08
27	141	4.5	351	14	O57117	TUMOR NECROSIS FACTOR	3.52e-08
28	141	4.5	368	13	O57408	SUBGROUP E ALV RECEPTO	3.52e-08
29	137	4.4	267	6	O02764	OX40 PRECURSOR (FRAGME	1.87e-07
30	138	4.4	322	14	O77761	K2R PROTEIN.	1.23e-07
31	137	4.4	347	14	O57119	TUMOR NECROSIS FACTOR	1.87e-07
32	137	4.4	351	14	O57121	TUMOR NECROSIS FACTOR	1.87e-07
33	137	4.4	314	14	O57118	TUMOR NECROSIS FACTOR	1.87e-07
34	133	4.3	160	14	O14293	FAS SOLUBLE PROTEIN.	9.72e-07
35	130	4.2	349	14	O57284	TUMOR NECROSIS FACTOR	3.29e-06
36	130	4.2	349	14	O57098	TUMOR NECROSIS FACTOR	3.29e-06
37	128	4.1	283	7	O92956	HERPESVIRUS ENTRY MEDI	6.13e-06
38	128	4.1	320	14	O57079	TUMOR NECROSIS FACTOR	7.37e-06
39	128	4.1	348	14	O84407	HOMOLOG OF VACCINIA VI	7.37e-06
40	128	4.1	348	14	O57112	TUMOR NECROSIS FACTOR	7.37e-06
41	128	4.1	349	14	O57110	TUMOR NECROSIS FACTOR	7.37e-06
42	128	4.1	349	14	O89118	SOMALIA-1977 RIGHT NEA	7.37e-06
43	128	4.1	349	14	O57111	TUMOR NECROSIS FACTOR	7.37e-06
44	128	4.1	349	14	O89098	GARCIA-1966 RIGHT NEAR	7.37e-06
45	128	4.1	471	6	O19131	TUMOR NECROSIS FACTOR-	7.37e-06

ALIGNMENTS

RESULT 1

```
ID   O72735;     PRELIMINARY;       PRT;   186 AA.
AC   O72735;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   A53R PROTEIN.
GN   A53R.
OS   Cowpox virus (CPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GRI-90;
RX   MEDLINE; 98229462.
RA   SHCHELKUNOV S.N., SAFRONOV P.F., TOTMENIN A.V., PETROV N.A.,
RA   RYAZANKINA O.I., GUTOROV V.V., KOTWAL G.J.;
RT   "The genomic sequence analysis of the left and right species-specific
RT   terminal region of a cowpox virus strain reveals unique sequences and
RT   a cluster of intact ORFs for immunomodulatory and host range
RT   proteins.";
RL   Virology 243:432-460(1998).
DR   EMBL; Y15035; CAA75273.1; -.
DR   HSSP; P25942; 1CDF.
DR   PFAM; PF00020; TNFR_c6; 2.
SQ   SEQUENCE   186 AA;   20482 MW;   C635B516 CRC32;

Query Match  6.6%;  Score 205;  DB 14;  Length 186;
Best Local Similarity  33.3%;
Matches   37;  Conservative   23;  Mismatches   40;  Indels   11;  Gaps   10;

Qy   34 NGSCDEGEYLDKRHNQC--CNQCPG-EFAKVRCS-G-SDNTKCRCPHTYTAIPNYSN 88
         :||: |||| | |:|||  |:| :: : || ||  ||:|:||  ::||  |: |:||:
Db   89 -CKPCLDCAVVNRFQKANCSATSDAICGDCLPGFYRKTKLVGFQD-MECVP 137

Qy   31 SGDCRQQEFHD-RSGNCVPCNQCGPGMELSKE-CGFGYGEDAQCVACRLHRFKEDWGFQK 88
Db   89 GCHQCRKCP'-GSFDKVKCIGTQNSKCS-CLPGWYCATDSSQCTEDCRDCVP 137
```

RESULT 2

```
ID   Q9WJB4;     PRELIMINARY;       PRT;   186 AA.
AC   Q9WJB4;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
```

```
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TUMOUR NECROSIS FACTOR RECEPTOR PRECURSOR.
GN   A53R.
OS   Vaccinia virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LISTER;
RX   MEDLINE; 99226947.
RA   ALCAMI A., KHANNA A., PAUL N., SMITH J.L.;
RT   "Vaccinia virus strains Lister, USSR and Evans express soluble and
RT   cell-surface tumour necrosis factor receptors.";
RL   J. Gen. Virol. 80:949-959(1999).
DR   EMBL; Y17728; CAB41042.1; -.
KW   Signal; Receptor.
FT   SIGNAL        1     22
SQ   SEQUENCE   186 AA;   20646 MW;   6D548A28 CRC32;

Query Match             6.5%;  Score 201;  DB 14;  Length 186;
Best Local Similarity   32.4%;
Matches  36;  Conservative  22;  Mismatches  42;  Indels  11;  Gaps  10;

Db   34  NGACDEGEYLDKRHNQC--CNOCPPG-EFAKYRCN-G-NDNTKCERCPPHTYTAIPNYSN 88
             :|:|: |  |::| |  |:||| |||:| |  |:|: :|| ::||::||   ::
Qy   31  SGDCRQQEFRD-RSGNGVPCNQCGPGMELSKE-CGFGYGEDAQCVACRLHRFKEDWGFQK 88

Db   89  GCHQCRKCPT-GSFDKVKCTGTQNSKCS-CLPGWYCATDSSQTEDCRDCIP 137
          |: ||:|   |||::| :|||::|  |  ||:|||:::::|::||:||:
Qy   89  -CKPCLDCAAVNRFQKANCSATSDAICGDCLPGFYRKTKLVGFQD-MECVP 137
```

```
RESULT  3
ID   Q9YP87
AC   Q9YP87;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   SOLUBLE TNF RECEPTOR CRMC.
GN   CRMC.
OS   Cowpox virus (CPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRIGHTON REDGENE;
RA   SMITH C.A., GOODWIN R.G., PICKUP D.J.;
RT   "Cowpox virus Encodes a second soluble TNF receptor (CrmC) Distinct
RT   from CrmB.";
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U55052; AAD10325.1; -.
KW   Receptor.
SQ   SEQUENCE   186 AA;   2A90DC98 CRC32;

Query Match             6.0%;  Score 187;  DB 14;  Length 186;
Best Local Similarity   30.0%;
Matches  33;  Conservative  28;  Mismatches  40;  Indels  9;  Gaps  9;

Db   34  NGSCDDGEYLDKTHN-QCCNRCPPG-EFAKIRCS-G-SDNTKCERCPPHTYTTVPNYSNG 89
             |::|   ||:  |  : ||:||| |||: ||: :  :|:|  :||  :  :
Qy   31  SGDCRQQEFRDRSGNGVPCNQCGPGMELSKE-CGFGYGEDAQCVACRLHRFKEDWGFQK- 88

Db   90  CHQCRKCPT-GSFDKVKCTGTQNSKCS-CLPGWFCATDSSKTEDCRDCIP 137
          |: ||:|   |||::| :|||::|  |  ||:||:::::|::||:||:
Qy   89  CKPCLDCAAVNRFQKANCSATSDAICGDCLPGFYRKTKLVGFQD-MECVP 137
```

```
RESULT  4
ID   Q9XS28
AC   Q9XS28;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
```

```
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   HVEAS.
GN   HVEAS.
OS   Cercopithecus aethiops (Green monkey) (Grivet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae;
OC   Chlorocebus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE; 99296730.
RA   FOSTER T.P., CHOULJENKO V.N., KOUSOULAS K.G.;
RT   "Functional characterization of the HveA homolog specified by African
RT   green monkey kidney cells with a herpes simplex virus expressing the
RT   green fluorescence protein.";
RL   Virology 258:365-374(1999).
DR   EMBL; AF147720; AAD37381.1; -.
DR   PROSITE; PS00652; TNFR_NGFR_1; 1.
SQ   SEQUENCE   283 AA;   30199 MW;   9B499EAB CRC32;

Query Match             5.2%;  Score 161;  DB 6;  Length 283;
Best Local Similarity   27.4%;
Matches  31;  Conservative  27;  Mismatches  46;  Indels  9;  Gaps  8;

Db   20  ILTLVLYILTFLGSSCYAPALPSCKEDEYPV-GSEC--CPKCGPGFHVRQACG-E-QTGTV 74
            |:|::|::  | ::|  :|||:|   ::||  :|: || |||||   | | :|:|
Qy   13  FFTLIVLLGYLSCKV-TCESGSDCRQQEFRDRSGNGVPCNQCGPGMELSKECGFGYGEDAQ 71

Db   75  CEPCSPGTYIAHFNGLSKCLCQQCMCDPAMGLRTSRNCSTTANALCG-CSPGHF 126
          |||| |||:|  |  |:|| ||  ||  | : : | :|::  ||  || ||||
Qy   72  CVACRLHRFKEDW-GFQKCKPCLDCAAVNRFQKA-NCSATSDAICGDCLPGFY 122
```

```
RESULT  5
ID   O57108
AC   O57108;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN   CRMB.
OS   Monkeypox virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ZAIRE-1970 (CONGO-8);
RA   LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U88142; AAB94367.1; -.
DR   HSSP; P25942; 1CDF.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PFAM; PF00020; TNFR_c6; 2.
SQ   SEQUENCE   348 AA;   38212 MW;   28F1C77C CRC32;

Query Match             4.9%;  Score 152;  DB 14;  Length 348;
Best Local Similarity   30.9%;
Matches  29;  Conservative  20;  Mismatches  36;  Indels  9;  Gaps  8;

Db   29  NGKCKDNEYRSRN-LC--CLSCPPGTYASRLCD-S-KTNTQCTPCGSDTFTSHNNHLQAC 83
            ||| ::| |::  ||   ||:|||||| :| | :  |:|:|| |||: |: ::: ||
Qy   31  SGDCRQQEFRDRSGNGVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKEDWG-FQKC 89

Db   84  LSCNGRCDS-NQVETRSCNTTHNRIC-ECSPGYY 115
          |:|: ||: |::  :||::|: || ||| || |
Qy   90  KPCLD-CAAVNRFQKANCSATSDAICGDCLPGFY 122
```

```
RESULT  6
ID   O57103
AC   O57103;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
```

DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN   CRMB.
OS   Monkeypox virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ZAIRE-1979 (79-0005);
RA   LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U87847; AAB94364.1; -.
DR   HSSP; P25942; 1CDF.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PFAM; PF00020; TNFR_c6; 2.
SQ   SEQUENCE   348 AA;   38184 MW;   EF87B2D2 CRC32;

Query Match   4.9%;   Score 152;   DB 14;   Length 348;
Best Local Similarity   30.9%;   Pred. No. 3.19e-10;
Matches   29;   Conservative   20;   Mismatches   36;   Indels   9;   Gaps   8;

Db   29 NGKCKDNEYRSRN-LC--CLSCPPGTYASRLCD-S-KINTQCTPCGSDTFTSHNNHLQAC 83
Qy   31 SGDCRQEEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKEDWG-FQKC 89

Db   84 LSCNGRCDS-NQVETRSCNTTHNRIC-ECSPGYY 115
Qy   90 KPCLD-CAVVNRFQKANCSATSDAICGDCLPGFY 122

RESULT   7
ID   O57277   PRELIMINARY;   PRT;   348 AA.
AC   O57277;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN   CRMB.
OS   Monkeypox virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ZAIRE-1996 (96-17), ZAIRE-1996 (96-16);
RA   LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U88543; AAB94378.1; -.
DR   EMBL; U87841; AAB94358.1; -.
DR   HSSP; P25942; 1CDF.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PFAM; PF00020; TNFR_c6; 2.
SQ   SEQUENCE   348 AA;   38212 MW;   1AAABB69 CRC32;

Query Match   4.9%;   Score 152;   DB 14;   Length 348;
Best Local Similarity   30.9%;   Pred. No. 3.19e-10;
Matches   29;   Conservative   20;   Mismatches   36;   Indels   9;   Gaps   8;

Db   29 NGKCKDNEYRSRN-LC--CLSCPPGTYASRLCD-S-KINTQCTPCGSDTFTSHNNHLQAC 83
Qy   31 SGDCRQEEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKEDWG-FQKC 89

Db   84 LSCNGRCDS-NQVETRSCNTTHNRIC-ECSPGYY 115
Qy   90 KPCLD-CAVVNRFQKANCSATSDAICGDCLPGFY 122

RESULT   8
ID   O57102   PRELIMINARY;   PRT;   349 AA.
AC   O57102;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.

GN   CRMB.
OS   Monkeypox virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.

RESULT   9
ID   O57101   PRELIMINARY;   PRT;   349 AA.
AC   O57101;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN   CRMB.
OS   Monkeypox virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BENIN-1978 (78-3945);
RA   LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U87846; AAB94363.1; -.
DR   HSSP; P25942; 1CDF.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PFAM; PF00020; TNFR_c6; 2.
SQ   SEQUENCE   349 AA;   38308 MW;   17182545 CRC32;

Query Match   4.9%;   Score 152;   DB 14;   Length 349;
Best Local Similarity   30.9%;   Pred. No. 3.19e-10;
Matches   29;   Conservative   20;   Mismatches   36;   Indels   9;   Gaps   8;

Db   29 NGKCKDNEYRSRN-LC--CLSCPPGTYASRLCD-S-KINTQCTPCGSDTFTSHNNHLQAC 83
Qy   31 SGDCRQEEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKEDWG-FQKC 89

Db   84 LSCNGRCDS-NQVETRSCNTTHNRIC-ECSPGYY 115
Qy   90 KPCLD-CAVVNRFQKANCSATSDAICGDCLPGFY 122

RESULT   10
ID   O57100   PRELIMINARY;   PRT;   349 AA.
AC   O57100;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN   CRMB.
OS   Monkeypox virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NIGERIA-1971 (71-0082);
RA   LOPAREV V. N., PARSONS J. M., ESPOSITO J. J.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U87844; AAB94361.1; -.
DR   HSSP; P25942; 1CDF.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PFAM; PF00020; TNFR_c6; 2.
SQ   SEQUENCE   349 AA; 38239 MW; F8871DD2 CRC32;

Query Match   4.9%; Score 152; DB 14; Length 349;
Best Local Similarity   30.9%; Pred. No. 3.19e-10;
  Matches   29; Conservative   20; Mismatches   36; Indels   9; Gaps   8;

Qy   29 NGKCKDNEYRSRN-LC--CLSCPPGTYASRLCD-S-KINTQCTPCGSDTFTSHNNHLQAC 83
Db   31 SGDCRQQEFRDRSGNCVPCNQCGPGMELSKECGFYGEDAQCVACRLHRFKEDWG-FQKC 89

Qy   84 LSCNGRCDS-NQVETRSCNTTHNRIC-ECSPGIYY 115
Db   90 KPCLD-CAVVNRFQKANCSATSDAICGDCLPGFY 122

RESULT 11

ID   057291   PRELIMINARY;   PRT;   349 AA.
AC   057291;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN   CRMB.
OS   Monkeypox virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VARIOUS STRAINS;
RA   LOPAREV V. N., PARSONS J. M., ESPOSITO J. J.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U88144; AAB94369.1; -.
DR   EMBL; U87842; AAB94359.1; -.
DR   EMBL; U87994; AAB94365.1; -.
DR   EMBL; U87995; AAB94366.1; -.
DR   EMBL; U88143; AAB94368.1; -.
DR   HSSP; P25942; 1CDF.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PFAM; PF00020; TNFR_c6; 2.
SQ   SEQUENCE   349 AA; 38295 MW; 7313FCF9 CRC32;

Query Match   4.9%; Score 152; DB 14; Length 349;
Best Local Similarity   30.9%; Pred. No. 3.19e-10;
  Matches   29; Conservative   20; Mismatches   36; Indels   9; Gaps   8;

Qy   29 NGKCKDNEYRSRN-LC--CLSCPPGTYASRLCD-S-KINTQCTPCGSDTFTSHNNHLQAC 83
Db   31 SGDCRQQEFRDRSGNCVPCNQCGPGMELSKECGFYGEDAQCVACRLHRFKEDWG-FQKC 89

Qy   84 LSCNGRCDS-NQVETRSCNTTHNRIC-ECSPGIYY 115
Db   90 KPCLD-CAVVNRFQKANCSATSDAICGDCLPGFY 122

RESULT 12

ID   057099   PRELIMINARY;   PRT;   349 AA.
AC   057099;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
DE   CRMB.
OS   Monkeypox virus.

OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SIERRA LEONE-1970 (70-0266);
RA   LOPAREV V. N., PARSONS J. M., ESPOSITO J. J.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U87843; AAB94360.1; -.
DR   HSSP; P25942; 1CDF.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PFAM; PF00020; TNFR_c6; 2.
SQ   SEQUENCE   349 AA; 38321 MW; 5315315C CRC32;

Query Match   4.9%; Score 152; DB 14; Length 349;
Best Local Similarity   30.9%; Pred. No. 3.19e-10;
  Matches   29; Conservative   20; Mismatches   36; Indels   9; Gaps   8;

Qy   29 NGKCKDNEYRSRN-LC--CLSCPPGTYASRLCD-S-KINTQCTPCGSDTFTSHNNHLQAC 83
Db   31 SGDCRQQEFRDRSGNCVPCNQCGPGMELSKECGFYGEDAQCVACRLHRFKEDWG-FQKC 89

Qy   84 LSCNGRCDS-NQVETRSCNTTHNRIC-ECSPGIYY 115
Db   90 KPCLD-CAVVNRFQKANCSATSDAICGDCLPGFY 122

RESULT 13

ID   000276   PRELIMINARY;   PRT;   253 AA.
AC   000276;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   LYMPHOCYTE ASSOCIATED RECEPTOR OF DEATH 2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97272273.
RA   SCREATON G.R., XU X.N., OLSEN A.L., COWPER A.E., TAN R.,
RA   MCMICHAEL A.J., BELL J.I.;
RT   "LARD: a new lymphoid-specific death domain containing receptor
RT   regulated by alternative pre-mRNA splicing.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:4615-4619(1997).
DR   EMBL; U94503; AAC51308.1; -.
DR   HSSP; P19438; 1TNR.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PFAM; PF00020; TNFR_c6; 2.
SQ   SEQUENCE   253 AA; 26934 MW; A21C863E CRC32;

Query Match   4.8%; Score 150; DB 4; Length 253;
Best Local Similarity   52.9%; Pred. No. 7.58e-10;
  Matches   18; Conservative   5; Mismatches   10; Indels   1; Gaps   1;

Qy   138 CQPCLDCGALHRHTRLLCS-RRDTDCGTCLPGFY 170
Db   89 CKPCLDCAVVNRFQKANCSATSDAICGDCLPGFY 122

RESULT 14

ID   014866   PRELIMINARY;   PRT;   277 AA.
AC   014866;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   SOLUBLE DEATH RECEPTOR 3 BETA.
DE   DR3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   WARZOCHA K., RIBEIRO P., RENARD N., CHARLOT C., COIFFIER B.,

```
RA   SALLES G.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF026071; AAB82288.1; -.
DR   HSSP; P19438; 1TNR.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PFAM; PF00020; TNFR_c6; 2.
SQ   SEQUENCE   277 AA;  29111 MW;  06E165C7 CRC32;

Query Match             4.8%;  Score 150;  DB 4;  Length 277;
Best Local Similarity   52.9%;
Matches   18;  Conservative    5;  Mismatches  10;  Indels   1;  Gaps   1;

Qy         89 CKPCLDCAVVNRFQKANCSATSDAICGDCLPGFY 122

Db        138 CQPCLDCGALHRHTRLLCS-RRDTDCGTCLPGFY 170
              | ||||||: ::|    ||  |:| || ||||||


RESULT    15
ID   O57116                PRELIMINARY;       PRT;    350 AA.
AC   O57116;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN   CRMB.
OS   Cowpox virus (CPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MUNICH OPV 89/2(CAT);
RA   LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U90226; AAB94382.1; -.
DR   HSSP; P25942; 1CDF.
DR   PFAM; PF00020; TNFR_c6; 2.
SQ   SEQUENCE   350 AA;  38216 MW;  5FF88E39 CRC32;

Query Match             4.8%;  Score 150;  DB 14;  Length 350;
Best Local Similarity   29.8%;
Matches   28;  Conservative   19;  Mismatches  40;  Indels   7;  Gaps   6;

Qy         29 NGKCKDNEY-NRHNLC--CLSCPPGTYASRLCDSKTNTNTQCTPCGSDTFTSRNNHLPAC 85

Db         31 SGDCRQEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKE-DWGFQKC 89
              :|:|::|  |: |  || |  |  || |  |:|| ||:  || |||| |||


Qy         90 KPCLD-CAVVNRFQKANCSATSDAICGDCLPGFY 122

Db         86 LSCNGRCDS-NQVETRSCNTHNRIC-DCAPGYY 117
              :|         |  :|:::|::: || ||| |:|


Search completed: Tue Apr 18 14:02:30 2000
Job time : 98 secs.
```
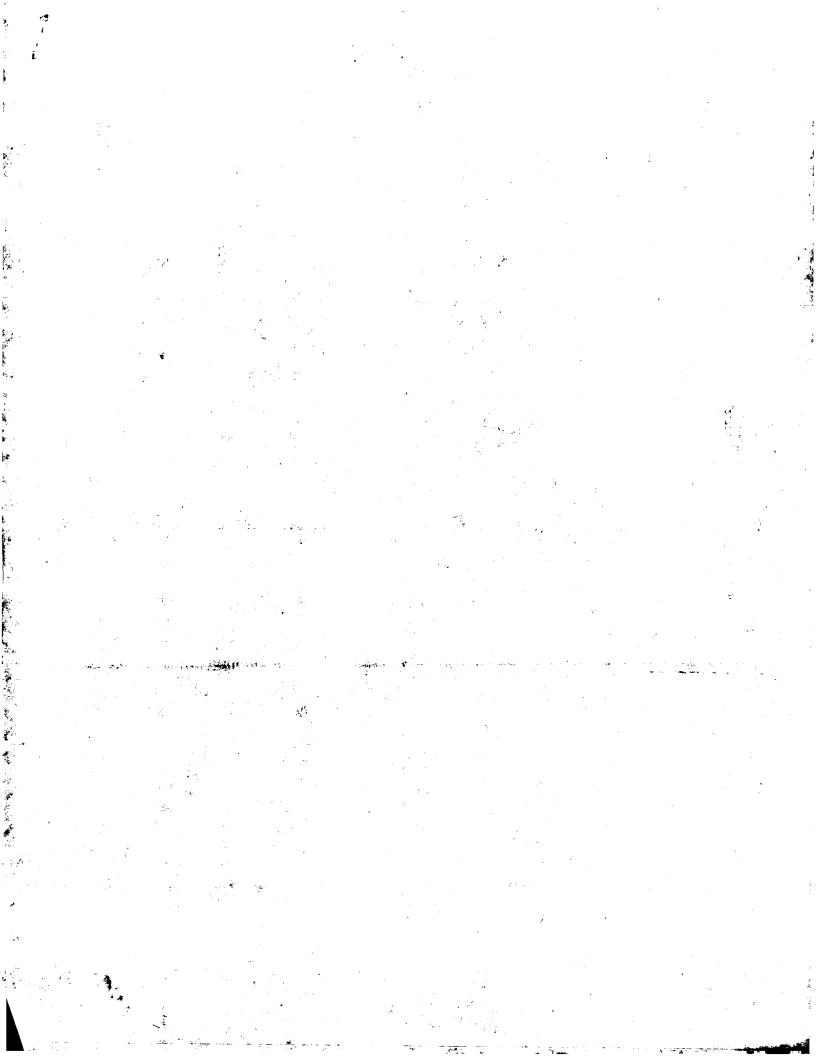
```
*********************************************************
*********************************************************
**                                                     **
**     M P s r c h                              (TM)    **
**                                                     **
*********************************************************
*********************************************************
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Apr 18 13:59:51 2000; MasPar time 39.25 Seconds
                                               321.889 Million cell updates/sec

Tabular output not generated.

Title:
Description: >US-09-490-187-2
                      (1-423) from US09490187.pep
Perfect Score: 3111
Sequence: 1 MALKVLLEQEKTFFTLLVLL..........AVIHPATQTSLQVRQRLGSL 423

Scoring table: PAM 150
                      Gap 11

Searched: 82229 seqs, 29864866 residues

Post-processing: Minimum Match 0%
                      Listing first 45 summaries

Database: swiss-prot38
                      1:swissprot

Statistics: Mean 48.063; Variance 74.860; scale 0.642

                      Pred. No. is the number of results predicted by chance to have a
                      score greater than or equal to the score of the result being printed,
                      and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	170	5.5	272	1	OX40_MOUSE	OX40L RECEPTOR PRECURS	1.9e-14
2	149	4.8	271	1	OX40_RAT	OX40L RECEPTOR PRECURS	2.86e-10
3	150	4.8	417	1	WSL1_HUMAN	WSL-1 PROTEIN PRECURSO	1.83e-10
4	146	4.7	416	1	NGFR_CHICK	LOW-AFFINITY NERVE GRO	1.08e-09
5	146	4.7	461	1	TNR1_PIG	TUMOR NECROSIS FACTOR	1.08e-09
6	142	4.6	454	1	TNR1_MOUSE	TUMOR NECROSIS FACTOR	6.22e-09
7	143	4.6	455	1	TNR1_HUMAN	TUMOR NECROSIS FACTOR	4.02e-09
8	136	4.4	435	1	TNRC_HUMAN	LYMPHOTOXIN-BETA RECEP	2.82e-08
9	133	4.3	335	1	FASL_HUMAN	FASL RECEPTOR PRECURSO	2.94e-07
10	130	4.2	326	1	VT2_MYXVL	TUMOR NECROSIS FACTOR	1.03e-06
11	130	4.2	415	1	TNRC_MOUSE	LYMPHOTOXIN-BETA RECEP	1.03e-06
12	130	4.2	461	1	TNR1_RAT	TUMOR NECROSIS FACTOR	1.03e-06
13	127	4.1	327	1	FASL_MOUSE	FASL RECEPTOR PRECURSO	3.58e-06
14	128	4.1	349	1	VC22_VARV	PROTEIN C22/B28 HOMOLO	2.37e-06
15	126	4.1	1192	1	LMG2_MOUSE	LAMININ GAMMA-2 CHAIN	5.39e-06
16	125	4.0	255	1	41BB_HUMAN	4-1BB LIGAND RECEPTOR	8.11e-06
17	125	4.0	474	1	TNR2_MOUSE	TUMOR NECROSIS FACTOR	8.11e-06
18	122	3.9	324	1	FASL_RAT	FASL RECEPTOR PRECURSO	2.73e-05
19	122	3.9	332	1	FASL_PIG	FASL RECEPTOR PRECURSO	2.73e-05
20	122	3.9	425	1	NGFR_RAT	LOW-AFFINITY NERVE GRO	2.73e-05
21	121	3.9	427	1	NGFR_HUMAN	LOW-AFFINITY NERVE GRO	4.08e-05
22	120	3.9	461	1	TNR2_HUMAN	TUMOR NECROSIS FACTOR	6.08e-05
23	119	3.8	323	1	FASA_BOVIN	FASL RECEPTOR PRECURSO	9.05e-05
24	118	3.8	1193	1	LMG2_HUMAN	LAMININ GAMMA-2 CHAIN	1.34e-04
25	115	3.7	1680	1	FUR2_DROME	FURIN-LIKE PROTEASE 2	4.34e-04
26	111	3.6	277	1	CD40_HUMAN	CD40L RECEPTOR PRECURS	2.01e-03
27	110	3.5	3075	1	LMA1_HUMAN	LAMININ ALPHA-1 CHAIN	2.94e-03
28	108	3.5	3084	1	LMA1_MOUSE	LAMININ ALPHA-1 CHAIN	6.21e-03
29	107	3.4	250	1	CD27_MOUSE	CD27L RECEPTOR PRECURS	9.00e-03
30	105	3.4	256	1	CD27_HUMAN	CD27L RECEPTOR PRECURS	1.88e-02
31	105	3.4	260	1	41BB_MOUSE	4-1BB LIGAND RECEPTOR	1.88e-02
32	106	3.4	277	1	OX40_HUMAN	OX40L RECEPTOR PRECURS	1.30e-02
33	106	3.4	703	1	SAT1_RAT	SULFATE ANION TRANSPOR	1.30e-02
34	106	3.4	2871	1	FBN1_MOUSE	FIBRILLIN 1 PRECURSOR.	1.30e-02
35	104	3.3	289	1	CD40_MOUSE	CD40L RECEPTOR PRECURS	2.70e-02
36	104	3.3	956	1	TSP3_MOUSE	THROMBOSPONDIN 3 PRECU	2.70e-02
37	103	3.3	2871	1	FBN1_BOVIN	FIBRILLIN 1 PRECURSOR	3.87e-02
38	103	3.3	3635	1	LMA5_MOUSE	LAMININ ALPHA-5 CHAIN	3.87e-02
39	99	3.2	103	1	VA53_VACCC	PROTEIN A53.	1.59e-01
40	99	3.2	103	1	VA53_VACCV	PROTEIN A53.	1.59e-01
41	99	3.2	215	1	OS25_PLAGA	25 KD OOKINETE SURFACE	1.59e-01
42	98	3.2	400	1	LMB_HIRME	LAMININ B-CHAIN (FRAGM	2.26e-01
43	99	3.2	1184	1	FBL2_HUMAN	FIBULIN-2 PRECURSOR.	1.59e-01
44	101	3.2	1607	1	LMG1_MOUSE	LAMININ GAMMA-1 CHAIN	7.90e-02
45	99	3.2	3106	1	LMA2_MOUSE	LAMININ ALPHA-2 CHAIN	1.59e-01

ALIGNMENTS

RESULT   1

ID   OX40_MOUSE   STANDARD;   PRT;   272 AA.
AC   P47741;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN).
GN   TXGP1 OR OX40.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C;
RX   MEDLINE; 94044750.
RA   CALDERHEAD D.M., BUHLMANN J.E., VAN DEN EERTWEGH A.J.,
RA   CLAASSEN E., NOELLE R.J., FELL H.;
RT   "Cloning of mouse Ox40: a T cell activation marker that may mediate
RT   T-B cell interactions.";
RL   J. Immunol. 15:5261-5271(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95255413.
RA   BIRKELAND M.L., COPELAND N.G., GILBERT D.J., JENKINS N.A.,
RA   BARCLAY A.N.;
RT   "Gene structure and chromosomal localization of the mouse homologue
RT   of rat OX40 protein.";
RL   Eur. J. Immunol. 25:926-930(1995).
RN   [2]
RL   J. Immunol. 15:5261-5271(1993).

CC   -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; Z21674; CAA79772.1; -.
DR   EMBL; X85214; CAA59476.1; -.
DR   HSSP; P25942; 1CDF.
DR   MGD; MGI:104512; TXGP1.
DR   PROSITE; PS00652; TNFR_NGFR_1; 3.
DR   PROSITE; PS50050; TNFR_NGFR_2; 2.
DR   PFAM; PF00020; TNFR_c6; 3.
```

```
RESULT 2

ID   OX40_RAT       STANDARD;      PRT;   271 AA.
AC   P15725;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN) (MRC OX40).
GN   TXGP1L OR OX40.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 90214614.
RA   MALLETT S., FOSSUM S., BARCLAY A.N.;
RT   "Characterization of the MRC OX40 antigen of activated CD4 positive T
RT   lymphocytes -- a molecule related to nerve growth factor receptor.";
RL   EMBO J. 9:1063-1068(1990).
CC   -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: ACTIVATED T-CELLS.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; X17037; CAA34897.1; -.
DR   PIR; S08036; S08036.
DR   PIR; S12783; S12783.
DR   HSSP; P25942; 1CDF.
DR   PROSITE; PS00652; TNFR_NGFR_1; 3.
DR   PROSITE; PS50050; TNFR_NGFR_2; 2.
DR   PFAM; PF00020; TNFR_c6; 3.
KW   Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
KW   Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    272       OX40L RECEPTOR.
FT   DOMAIN       20    211       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    212    236       POTENTIAL.
FT   DOMAIN      237    272       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      165    272       4 X TNFR-CYS.
FT   REPEAT       26     61       TNFR-CYS 1.
FT   REPEAT       62    103       TNFR-CYS 2.
FT   REPEAT      104    124       TNFR-CYS 3 (INCOMPLETE).
FT   REPEAT      125    165       TNFR-CYS 4.
FT   CARBOHYD    144    144       POTENTIAL.
FT   CARBOHYD     15     15       POTENTIAL.
FT   CONFLICT     15     15       A -> G (IN REF. 2).
SQ   SEQUENCE   272 AA;  30153 MW;  67D1B978 CRC32;

Query Match             5.5%;  Score 170;  DB 1;  Length 272;
Best Local Similarity  27.3%;  Pred. No. 1.94e-14;
Matches   48;  Conservative   36;  Mismatches  78;  Indels  14;  Gaps  11;

Qy    9  TALILLA-LILGVTARRLNCVKHTYPS-GHKC--CRECQPGHGMVSRCD--HTRDTLCHP  62
            :||:||||  :|||| |||            |||  |||  |  |    :||
Db   15  TLLVLLGYLSCKVTCESGDCRQEEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVA  74

Qy   63  CETGFYNEAVNYDTCKQCTQCNHRSGSELKQNCTPTQDTVC-RCRPGTQPRQD-SGYKLG 120
            ::|   |||:|||||                 |  ||| ||||    :|
Db   75  CRLHRFKEDWGFQKCKPCLDCAVVNRFQ-KANCSATSDAICGDCLPGFYRKTKLVGFQ-D 132

Qy  121  VDCVPCPPGHFSPGNNQACKPWTNCTLSGKQTRHPASDSLDAV-CED-RSLLATLL     174
           ||||      ::|                                |
Db  133  MECVPC--GDPPPYEPHCASKVNLVKIASTASSPRDIALAAVICSALATVLLALL     186
```

```
RESULT 3

ID   WSL1_HUMAN     STANDARD;      PRT;   417 AA.
AC   Q93038; Q93036; Q93037; Q92983; P78515; Q99831; Q99722; P78507;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-DEC-1999 (Rel. 39, Last annotation update)
DE   WSL-1 PROTEIN PRECURSOR (APOPTOSIS-MEDIATING RECEPTOR DR3) (APOPTOSIS-
DE   MEDIATING RECEPTOR TRAMP) (DEATH DOMAIN RECEPTOR 3) (WSL PROTEIN)
DE   (APOPTOSIS INDUCING RECEPTOR AIR) (APO-3) (LYMPHOCYTE ASSOCIATED
DE   RECEPTOR OF DEATH) (LARD).
GN   TNFRSF12 OR WSL1 OR WSL OR APO3 OR DR3 OR DDR3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND MUTAGENESIS.
RC   TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;
RX   MEDLINE; 97081063.
RA   CHINNAIYAN A.M., O'ROURKE K., YU G.-L., LYONS R.H., GARG M.,
RA   DUAN D.R., XING L., GENTZ R., NI J., DIXIT V.M.;
RT   "Signal transduction by DR3, a death domain-containing receptor
RT   related to TNFR-1 and CD95.";
RL   Science 274:990-992(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LYMPHOID;
RX   MEDLINE; 97088617.
RA   KITSON J., RAVEN T., JIANG Y.-P., GOEDDEL D.V., GILES K.M., PUN K.-T.,
RA   GRINHAM C.J., BROWN R., FARROW S.N.;
RT   "A death domain-containing receptor that mediates apoptosis.";
RL   Nature 384:372-375(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEART;
RA   MARSTERS S.A., SHERIDAN J.P., DONAHUE C.J., PITTI R.M., GRAY C.L.,
RA   GODDARD A.D., BAUER K.D., ASHKENAZI A.;
RT   "Apo-3, a new member of the tumor necrosis factor receptor family,
RT   contains a death domain and activates apoptosis and NF-kappa-B.;
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LYMPHOID;
RA   DEGLI-ESPOSTI M.A., DIN W.S., COSMAN D., SMITH C.A., GOODWIN R.G.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
KW   Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
KW   Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    271       OX40L RECEPTOR.
FT   DOMAIN       20    210       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    211    235       POTENTIAL.
FT   DOMAIN      236    271       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      164    271       4 X TNFR-CYS.
FT   REPEAT       25     60       TNFR-CYS 1.
FT   REPEAT       61    102       TNFR-CYS 2.
FT   REPEAT      103    123       TNFR-CYS 3 (INCOMPLETE).
FT   REPEAT      124    164       TNFR-CYS 4.
FT   CARBOHYD    143    143       POTENTIAL.
SQ   SEQUENCE   271 AA;  29895 MW;  65460ED2 CRC32;

Query Match             4.8%;  Score 149;  DB 1;  Length 271;
Best Local Similarity  29.1%;  Pred. No. 2.86e-10;
Matches   39;  Conservative   27;  Mismatches  57;  Indels  11;  Gaps   8;

Qy    9  TAFLLIG-LSLGVTVKL-NCVKDTYPS-GHKC--CRECQPGHGMVSRCD--HTRDTVCHP  61
Db   15  TLLVLLGYLSCKVTCESGDCRQEEFRDRSGNCVPCNQCGPGMELSKECGFGYGDSSHKLGV 74

Qy   62  CEPGFYNEAVNYDTCKQCTQCNHRSGSELKQNCTPTEDTVC-QCRPGTQPRQDSSHKLGV 120
Db   75  CRLHRFKEDWGFQKCKPCLDCAVVNRFQ-KANCSATSDAICGDCLPGFYRKTKLVGFQDM 133

Qy  121  DCVPC--PPGHFSP                                              132
Db  134  ECVPCGDPPPYEP                                              147
```

```
RL   Curr. Biol. 6:1669-1676(1996).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   SCREATON G.R., XU X.-N., OLSEN A.L., COMPER A.E., TAN R.,
RA   MCMICHAEL A.J., BELL J.I.;
RT   "LARD: a new lymphoid-specific death domain containing receptor
RT   regulated by alternative pre-mRNA splicing.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:4615-4619(1997).
RN   [6]
RP   SEQUENCE OF 4-417 FROM N.A.
RC   TISSUE=BRAIN, AND FETAL LUNG;
RX   MEDLINE; 97205335.
RA   BODMER J.-L., BURNS K., SCHNEIDER P., HOFMANN K., STEINER V.,
RA   THOME M., BORNAND T., HAHNE M., SCHROETER M., WILSON A., FRENCH L.E.,
RA   BROWNING J.L., MACDONALD H.R., TSCHOPP J.;
RT   "TRAMP, a novel apoptosis-mediating receptor with sequence homology
RT   to tumor necrosis factor receptor 1 and Fas(Apo-1/CD95).";
RL   Immunity 6:79-88(1997).
RN   [7]
RP   SEQUENCE OF 7-417 FROM N.A.
RC   TISSUE=BRAIN;
RA   CHAUDHARY P.M., HOOD L.E.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INDUCES APOPTOSIS AND ACTIVATES NUCLEAR FACTOR KAPPA-B
CC       (NF-KAPPAB). DIRECTLY INTERACTS WITH TRADD ADAPTATOR MOLECULE. MAY
CC       PLAY A ROLE IN REGULATING LYMPHOCYTE HOMEOSTASIS.
CC   -!- SUBUNIT: HOMODIMER. INTERACTS STRONGLY VIA THE DEATH DOMAINS WITH
CC       THE TNFR1-ASSOCIATED MOLECULE TRADD AND THE TNFR1 RECEPTOR TO
CC       ACTIVATE AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC       NF-KAPPA B SIGNALING.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC   -!- ALTERNATIVE PRODUCTS: THREE ISOFORMS; WSL-1/LARD-1A (SHOWN HERE),
CC       WSL-S1/LARD-3 AND WSL-S2; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN THYMOCYTES AND
CC       LYMPHOCYTES. DETECTED IN LYMPHOCYTE-RICH TISSUES SUCH AS THYMUS,
CC       COLON, INTESTINE, AND SPLEEN. ALSO FOUND IN THE PROSTATE.
CC   -!- PTM: GLYCOSYLATED (PROBABLE).
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y09392; CAA70561.1; -.
DR   EMBL; Y09392; CAA70559.1; -.
DR   EMBL; Y09392; CAA70560.1; -.
DR   EMBL; U72763; AAC50819.1; -.
DR   EMBL; U83599; AAB41434.1; -.
DR   EMBL; U83600; AAB41435.1; -.
DR   EMBL; U78029; AAB40918.1; -.
DR   EMBL; U74611; AAB39714.1; -.
DR   EMBL; U94501; AAC51306.1; -.
DR   EMBL; U94504; AAC51309.1; -.
DR   EMBL; U75380; AAC51192.1; -.
DR   EMBL; U75381; AAC51193.1; -.
DR   EMBL; U83597; AAB41432.1; -.
DR   HSSP; P19438; 1TNR.
DR   MIM; 603366;
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PROSITE; PS50050; TNFR_NGFR_2; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PFAM; PF00020; TNFR_c6; 2.
DR   PFAM; PF00531; death; 1.
KW   Receptor; Apoptosis; Alternative splicing; Transmembrane; Signal;
KW   Repeat.
FT   SIGNAL          1     24       POTENTIAL.
FT   CHAIN          25    417       WSL-1 PROTEIN.
FT   DOMAIN        200    199       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      200    220       POTENTIAL.
FT   DOMAIN        221    417       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN         34    192       4 X TNFR-CYS.
FT   REPEAT         34     71       TNFR-CYS 1.
FT   REPEAT         72    115       TNFR-CYS 2.
FT   REPEAT        115    163       TNFR-CYS 3.
FT   REPEAT        161    192       TNFR-CYS 4.
FT   DOMAIN        332    413       DEATH DOMAIN.
FT   CARBOHYD        6?     6?      POTENTIAL.
FT   CARBOHYD      106    106       POTENTIAL.
FT   CARBOHYD      182    218       POTENTIAL.
FT   VARSPLIC      219    417       STLGSCPERCAAVCGWRQMFWVQVLLAGLVPLLLGA ->
FT                                  VLGPGAPGWPGCPPPAWGHPDLHIPPLLASQAPGYCR (IN
FT                                  ISOFORM WSL-S1).
FT   VARSPLIC      254    417       MISSING (IN ISOFORM WSL-S1).
FT   VARSPLIC      200    253       MFWVQVLLAGLVVPLLLGATILIYTYRHCWPHKPLVTADEAG
FT                                  MEALITPPPATHLS -> SRWCAGNARGRTGMDRGEAGEEGG
FT                                  NHPTPTSCFQCGSGSRCSWLAIWSPSCLGPP (IN
FT                                  ISOFORM WSL-S2).
FT   VARSPLIC      254    417       MISSING (IN ISOFORM WSL-S2).
FT   MUTAGEN       354    354       L->A: SUPPRESSES HOMODIMERIZATION, TNFR1
FT                                  INTERACTION, AND APOPTOSIS INDUCTION.
FT   MUTAGEN       356    356       L->A: SUPPRESSES HOMODIMERIZATION, AND
FT                                  TNFR1 INTERACTION.
FT   MUTAGEN       373    373       D->A: SUPPRESSES HOMODIMERIZATION, AND
FT                                  TNFR1 INTERACTION.
FT   CONFLICT        4      6       RPR -> AAA (IN REF. 6).
FT   CONFLICT       60     60       P -> H (IN REF. 7).
FT   CONFLICT      167    167       P -> L (IN REF. 6 AND 7).
FT   CONFLICT      312    312       A -> R (IN REF. 1).
FT   CONFLICT      370    370       R -> L (IN REF. 1).
FT   CONFLICT      381    381       R -> H (IN REF. 1).
SQ   SEQUENCE   417 AA;  45385 MW;  3AEEF0A5 CRC32;
```

```
Query Match           4.8%
Best Local Similarity  52.9%;
Matches     18;   Conservative   5;   Mismatches   10;   Indels   1;   Gaps   1;

Score   150;   DB 1;   Length 417;
Pred. No. 1.83e-10;

Db   138  COPCLDCGALRHRTRLLCS-RRDTDCGTCLPGFY 170
          | |||||:  ::| |  ||  || |||| |||||
Qy    89  CKPCLDCAVVNRFQKANCSATSDAICGDCLPGFY 122
```

```
RESULT       4
ID    NGFR_CHICK      STANDARD;      PRT;      416 AA.
AC    P18519;
DT    01-NOV-1990 (Rel. 16, Created)
DT    01-NOV-1990 (Rel. 16, Last sequence update)
DT    15-JUL-1999 (Rel. 38, Last annotation update)
DE    LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR)
DE    (GP80-LNGFR) (P75 ICD).
GN    NGFR.
OS    Gallus gallus (Chicken).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC    Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=BRAIN;
RX    MEDLINE; 90166579.
RA    LARGE T.H., WESKAMP G., HELDER J.C., RADEKE M.J., MISKO T.P.,
RA    SHOOTER E.M., REICHARDT L.F.;
RT    "Structure and developmental expression of the nerve growth factor
RT    receptor in the chicken central nervous system.";
RL    Neuron 2:1123-1134(1989).
RN    [2]
RP    SEQUENCE OF 21-416 FROM N.A.
RX    MEDLINE; 90152140.
RA    HEUER J.G., FATEMIE-NAINIE S., WHEELER E.F., BOTHWELL M.;
RT    "Structure and developmental expression of the chicken NGF receptor.";
RL    Dev. Biol. 137:287-304(1990).
CC    -!- FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,
CC        NT-3, AND NT-4.
CC    -!- SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE
CC        BOND FORMATION.
CC    -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
```

```
CC   -!- PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR   PIR; JN0006; JN0006.
DR   PIR; A60504; A60504.
     HSSP; P07174; 1NGR.
DR   PROSITE; PS00652; TNFR_NGFR_1; 3.
DR   PROSITE; PS00050; TNFR_NGFR_2; 3.
DR   PROSITE; PS00017; DEATH_DOMAIN; 1.
     PFAM; PF00020; TNFR_c6; 4.
     PFAM; PF00531; death; 1.
KW   Receptor; Neurogenesis; Transmembrane; Glycoprotein; Repeat;
KW   Phosphorylation; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    416   LOW-AFFINITY NERVE GROWTH FACTOR
FT                            RECEPTOR.
FT   DOMAIN       29    239   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    240    261   POTENTIAL.
FT   DOMAIN      262    416   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       23    181   4 X TNFR-CYS.
FT   REPEAT       23     57   TNFR-CYS 1.
FT   REPEAT       58    100   TNFR-CYS 2.
FT   REPEAT      101    140   TNFR-CYS 3.
FT   REPEAT      141    181   TNFR-CYS 4.
FT   DOMAIN      188    236   SER/THR-RICH.
FT   DOMAIN      333    410   DEATH DOMAIN.
FT   DISULFID     24     35   BY SIMILARITY.
FT   DISULFID     36     49   BY SIMILARITY.
FT   DISULFID     39     56   BY SIMILARITY.
FT   DISULFID     59     75   BY SIMILARITY.
FT   DISULFID     78     91   BY SIMILARITY.
FT   DISULFID     81     99   BY SIMILARITY.
FT   DISULFID    101    114   BY SIMILARITY.
FT   DISULFID    117    130   BY SIMILARITY.
FT   DISULFID    120    138   BY SIMILARITY.
FT   DISULFID    141    156   BY SIMILARITY.
FT   DISULFID    159    172   BY SIMILARITY.
FT   DISULFID    162    180   BY SIMILARITY.
FT   CARBOHYD     52     52   POTENTIAL.
FT   CONFLICT     36     36   C -> Y (IN REF. 2).
FT   CONFLICT    173    173   T -> K (IN REF. 2).
FT   CONFLICT    276    276   N -> S (IN REF. 2).
FT   CONFLICT    396    396   K -> R (IN REF. 2).
SQ   SEQUENCE    416 AA;  44654 MW;  4D3F086A CRC32;

Query Match    4.7%;  Score 146;  DB 1;  Length 416;
Best Local Similarity   30.0%;  Pred. No. 1.08e-09;
Matches   54;  Conservative   29;  Mismatches   83;  Indels   14;  Gaps   13;

Db    90  PC-VESDDAVCRCAYGYFQDELSGSCKECSICEVGFGLMFPCR-D-SQDTVCEECPEGTF  146
              ||:|      |:|||| |||    ||:|  |   ||:| |  | :||  |||  |
Qy    24  SCKVTCESGDCR-QQE-FRDR-SGNCVPCNCQCGPGMELSKECGFGYGEDAQCVACRLHRF  80

Db   147  SDEANFVDPCLPCTICEENEVWVK-ECTATSDAECRDLHPRWTHTPTSLAGSDSPEPITR  205
               | ||  |  || |||  | ::||  || :|| |   |    | |||  |:|||
Qy    81  KEDWGF-QKCRPCLDCAVVNRFQKANCSATSDAICGDCLPGFYRKI-KLVGFQDMECVPC  138

Db   206  DFFNTEGMATTLADIVTTVMGSSOPVVSRGTRADNLIPYYCSILAAVVVGL-VAYIAF-KR  263
              :     |: :: :|  ||  |:|| |:  ||  |||||  | ||  |||||:||||||
Qy   139  GD-PPPYEPHCASKVNLVKIASTASSPRDTA--LAAVICSALATVLLALLILCVIICKR  195

RESULT   5
ID   TNR1_PIG      STANDARD;      PRT;      461 AA.
AC   P50555;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-R1) (P55).
GN   TNFRSF1A OR TNFR1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE; 96011645.
RA   SUTER B., PAULI U.H.;
RT   "Cloning of the cDNA encoding the porcine p55 tumor necrosis factor
RT   receptor.";
RL   Gene 163:263-266(1995).
CC   -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
CC       RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
CC       AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
CC       PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
CC       SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE
CC       PROTEASES) MEDIATING APOPTOSIS.
CC   -!- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC       HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
CC       PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
CC       WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
CC       PROTEINS SUCH AS TRAFS, RIP AND POSSIBLY FADD, ARE RECRUITED TO
CC       TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
CC       ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC       NF-KAPPA B SIGNALING (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
```
```
DR   EMBL; U19994; AAC48499.1; -.
DR   HSSP; P19438; 1TNR.
DR   PROSITE; PS00652; TNFR_NGFR_1; 3.
DR   PROSITE; PS00050; TNFR_NGFR_2; 2.
DR   PROSITE; PS00017; DEATH_DOMAIN; 1.
DR   PFAM; PF00020; TNFR_c6; 3.
     PFAM; PF00531; death; 1.
KW   Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Apoptosis.
FT   SIGNAL        1     21
FT   CHAIN        22    461   TUMOR NECROSIS FACTOR RECEPTOR 1.
FT   DOMAIN       22    210   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    211    233   POTENTIAL.
FT   DOMAIN      234    461   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       43    195   4 X TNFR-CYS.
FT   REPEAT       43     82   TNFR-CYS 1.
FT   REPEAT       83    125   TNFR-CYS 2.
FT   REPEAT      126    166   TNFR-CYS 3.
FT   REPEAT      167    195   TNFR-CYS 4.
FT   DOMAIN      340    350   N-SMASE ACTIVATION DOMAIN (NSD).
FT   DOMAIN      362    447   DEATH DOMAIN.
FT   DISULFID     44     58   BY SIMILARITY.
FT   DISULFID     59     72   BY SIMILARITY.
FT   DISULFID     62     81   BY SIMILARITY.
FT   DISULFID     84     99   BY SIMILARITY.
FT   DISULFID    102    117   BY SIMILARITY.
FT   DISULFID    105    125   BY SIMILARITY.
FT   DISULFID    127    143   BY SIMILARITY.
FT   DISULFID    146    158   BY SIMILARITY.
FT   DISULFID    149    166   BY SIMILARITY.
FT   DISULFID    168    179   BY SIMILARITY.
FT   DISULFID    182    190   BY SIMILARITY.
FT   DISULFID    185    194   BY SIMILARITY.
FT   CARBOHYD     54     54   POTENTIAL.
FT   CARBOHYD     86     86   POTENTIAL.
FT   CARBOHYD    145    145   POTENTIAL.
FT   CARBOHYD    151    151   POTENTIAL.
SQ   SEQUENCE    461 AA;  50696 MW;  8E2C350A CRC32;

Query Match    4.7%;  Score 146;  DB 1;  Length 461;
Best Local Similarity   25.5%;  Pred. No. 1.08e-09;
Matches   28;  Conservative   22;  Mismatches   55;  Indels   5;  Gaps   5;
```

```
Db      44  CPQGKYSHPQNRSICCTKCHKGTYLHNDC-LGPGLDTDCRECDNGTFTASENHLTQCLSC 102
                |:|    |  :|:  ::||:|:||| |   :|| |::|  | ||   :|  |  | :|
Qy      34  CRQQEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKE-DWGFQCKPC  92

Db     103  SKCRSEMSQVEISPCTVDRDIVCG-CRKNQYRKYWSETLFQCLNCSLCPN 151
                |::||: :||  |    :|:| |  |::|  | |  || ||| |  : |
Qy      93  LDC-AVVNRFQKANCSATSDAICGDCLPGFYRKT-KLVGFEQDMECYPCGD 140
```

```
RESULT   6

ID   TNR1_MOUSE    STANDARD;    PRT;    454 AA.
AC   P25118;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-R1) (P55).
GN   TNFRSF1A OR TNFR1 OR TNFR-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91187885.
RA   LEWIS M., TARTAGLIA L.A., LEE A., BENNETT G.L., RICE G.C.,
RA   WONG G.H., CHEN E.Y., GOEDDEL D.V.;
RT   "Cloning and expression of cDNAs for two distinct murine tumor
RT   necrosis factor receptors demonstrate one receptor is species
RT   specific.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91246168.
RA   GOODWIN R.G., ANDERSON D., JERZY R., DAVIS T., BRANNAN C.I.,
RA   COPELAND N.G., JENKINS N.A., SMITH C.A.;
RT   "Molecular cloning and expression of the type 1 and type 2 murine
RT   receptors for tumor necrosis factor.";
RL   Mol. Cell. Biol. 11:3020-3026(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91285014.
RA   GRAY P.W., BARRETT K., TAYLOR-FISHWICK D.A., COPE A.P., KISSONERGHIS A.M.,
RA   BROCKHAUS M., GENTZ R., LESSLAUER W.;
RT   "Cloning, expression and cross-linking analysis of the murine p55
RT   tumor necrosis factor receptor.";
RL   Eur. J. Immunol. 21:1649-1656(1991).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94245292.
RA   ROTHE J.G., FELDMANN M., FOXWELL B.M.J.;
RT   "Molecular cloning and expression of the TNF type I receptor from a mouse
RT   endothelioma cell line.";
RL   Immunogenetics 39:450-451(1994).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SPLEEN;
RX   MEDLINE; 92039815.
RA   BEBO B.F., LINTHICUM D.S.;
RT   "Nucleotide sequence of the mouse Tnf receptor type b.";
RL   Immunogenetics 34:338-340(1991).
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93156721.
RA   ROTHE J., BLUETHMANN H., GENTZ R., LESSLAUER W., STEINMETZ M.;
RT   "Genomic organization and promoter function of the murine tumor
RT   necrosis factor receptor beta gene.";
RL   Mol. Immunol. 30:165-175(1993).
CC   -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
CC       RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
CC       AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
CC       PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
CC       SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE
CC       PROTEASES) MEDIATING APOPTOSIS (BY SIMILARITY).
CC   -!- SUBUNIT: THE BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC       HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
CC       PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
CC       WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
CC       PROTEINS SUCH AS TRAFS, RIP AND POSSIBLY FADD, ARE RECRUITED TO
CC       TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
CC       ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC       NF-KAPPA B SIGNALING (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M60468; AAA39751.1; -.
DR   EMBL; M59377; AAA40464.1; -.
DR   EMBL; X59238; CAA41922.1; -.
DR   EMBL; X57796; CAA40936.1; -.
DR   EMBL; L26349; AAA59361.1; -.
DR   EMBL; M76656; AAA40465.1; -.
DR   EMBL; M80067; AAA40465.1; JOINED.
DR   EMBL; M76655; AAA40465.1; JOINED.
DR   EMBL; A38634; GOMST1.
DR   PIR; S16677; S16677.
DR   PIR; S19021; S19021.
DR   HSSP; P19438; 1EXT.
DR   MGD; MGI:1314884; TNFRSF1A.
DR   PROSITE; PS00652; TNFR_NGFR_1; 3.
DR   PROSITE; PS00050; TNFR_NGFR_2; 3.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PFAM; PF00020; TNFR_c6; 4.
DR   PFAM; PF00531; death; 1.
KW   Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Apoptosis.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    454       TUMOR NECROSIS FACTOR RECEPTOR 1.
FT   DOMAIN       22    212       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    213    235       POTENTIAL.
FT   DOMAIN      236    454       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       43    196       4 X TNFR-CYS.
FT   REPEAT       43     82       TNFR-CYS 1.
FT   REPEAT       83    125       TNFR-CYS 2.
FT   REPEAT      126    166       TNFR-CYS 3.
FT   REPEAT      167    196       TNFR-CYS 4.
FT   DOMAIN      339    441       DEATH DOMAIN.
FT   DOMAIN      356    441       N-SMASE ACTIVATION DOMAIN (NSD).
FT   DOMAIN       44     58       BY SIMILARITY.
FT   DISULFID     59     72       BY SIMILARITY.
FT   DISULFID     62     81       BY SIMILARITY.
FT   DISULFID     84     99       BY SIMILARITY.
FT   DISULFID    102    117       BY SIMILARITY.
FT   DISULFID    105    125       BY SIMILARITY.
FT   DISULFID    127    143       BY SIMILARITY.
FT   DISULFID    146    158       BY SIMILARITY.
FT   DISULFID    149    166       BY SIMILARITY.
FT   DISULFID    168    179       BY SIMILARITY.
FT   DISULFID    182    191       BY SIMILARITY.
FT   DISULFID    185    195       BY SIMILARITY.
FT   CARBOHYD     54     54       POTENTIAL.
FT   CARBOHYD    151    151       POTENTIAL.
FT   CARBOHYD    202    202       POTENTIAL.
FT   CONFLICT    394    394       R -> G (IN REF. 6).
SQ   SEQUENCE   454 AA;  50129 MW;  4B6EEC09 CRC32;
```

```
Query Match        4.6%;     Score 142;    DB 1;    Length 454;
Best Local Similarity   26.9%;
Matches    29;   Conservative   16;   Mismatches   58;   Indels   5;   Gaps   5;
```

```
Db     44  CPQGKYHSKXNNSICCITCHKGTYLLVSDCPSP-GRDTVCRECEKGTFTASQNYLRQCLSC 102
```

```
Qy    34 CRQQEFRDRSGNCVPCNQCGPGMELSKEGCGFGYGEDAQCVACLHRFKEDWGFQK-CKPC  92
         | |  | | |   | |:| |  |  :  |  |   | :  | : :|   |  | : |:|
Db   103 KTCRKEMSQVEISPCQADKDTVCG-CKENQFQRYLSETHFQCVDCSPC             149

Qy    93 LDCA-VVNRFQKANCSATSDAICGDCLPG-FYRKTKLVGFQDMECVPC             138
```

RESULT 7

```
ID   TNR1_HUMAN     STANDARD;      PRT;      455 AA.
AC   P19438;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (TUMOR NECROSIS FACTOR
DE   BINDING PROTEIN 1) (TBPI) (P60) (P55) (TNF-R1) (CD120A).
GN   TNFRSF1A OR TNFR1 OR TNFAR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE; 90235285.
RA   SCHALL T.J., LEWIS M., KOLLER K.J., LEE A., RICE G.C., WONG G.H.W.,
RA   GETANAGA T., GRANGER G.A., LENTZ R., RAAB H., KOHR W.J., GOEDDEL D.V.;
RT   "Molecular cloning and expression of a receptor for human tumor
RT   necrosis factor.";
RL   Cell 61:361-370(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 90235284.
RA   LOETSCHER H., PAN Y.-C.E., LAHM H.-W., GENTZ R., BROCKHAUS M.,
RA   TABUCHI H., LESSLAUER W.;
RT   "Molecular cloning and expression of the human 55 kd tumor necrosis
RT   factor receptor.";
RL   Cell 61:351-359(1990).
RN   [3]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 41-53; 110-124 AND 199-201.
RX   MEDLINE; 91006021.
RA   NOPHAR Y., KEMPER O., BRAKEBUSCH C., ENGELMANN H., ZWANG R.,
RA   ADERKA D., HOLTMANN H., WALLACH D.;
RT   "Soluble forms of tumor necrosis factor receptors (TNF-Rs). The cDNA
RT   for the type I TNF-R, cloned using amino acid sequence data of its
RT   soluble form, encodes both the cell surface and a soluble form of the
RT   receptor.";
RL   EMBO J. 9:3269-3278(1990).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91017509.
RA   GRAY P.W., BARRETT K., CHANTRY D., TURNER M., FELDMAN M.;
RT   "Cloning of human tumor necrosis factor (TNF) receptor cDNA and
RT   expression of recombinant soluble TNF-binding protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:7380-7384(1990).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92250049.
RA   FUCHS P., STREHL S., DWORZAK M., HIMMLER A., AMBROS P.F.;
RT   "Structure of the human TNF receptor 1 (p60) gene (TNFR1) and
RT   localization to chromosome 12p13.";
RL   Genomics 13:219-224(1992).
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91090841.
RA   HIMMLER A., MAURER-FOGY I., KRONKE M., SCHEURICH P., PFIZENMAIER K.,
RA   LANTZ M., OLSSON I., HAUPTMANN R., STRATOWA C., ADOLF G.R.;
RT   "Molecular cloning and expression of human and rat tumor necrosis
RT   factor receptor chain (p60) and its soluble derivative, tumor
RT   necrosis factor-binding protein.";
RL   DNA Cell Biol. 9:705-715(1990).
RN   [7]
RP   SEQUENCE OF 41-45.
```

```
RX   MEDLINE; 90110215.
RA   ENGELMANN H., NOVICK D., WALLACH D.;
RT   "Two tumor necrosis factor-binding proteins purified from human
RT   urine. Evidence for immunological cross-reactivity with cell surface
RT   tumor necrosis factor receptors.";
RL   J. Biol. Chem. 265:1531-1536(1990).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.85 ANGSTROMS) OF 30-211.
RX   MEDLINE; 93258809.
RA   BANNER D.W., D'ARCY A., JANES W., GENTZ R., SCHOENFELD H.-J.,
RA   BROGER C., LOETSCHER H., LESSLAUER W.;
RT   "Crystal structure of the soluble human 55 kd TNF receptor-human TNF
RT   beta complex: implications for TNF receptor activation.";
RL   Cell 73:431-445(1993).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 41-202.
RX   MEDLINE; 97094982.
RA   NAISMITH J.H., DEVINE T.Q., KHONO H., SPRANG S.R.;
RT   "Structures of the extracellular domain of the type I tumor necrosis
RT   factor receptor.";
RL   Structure 4:1251-1262(1996).
CC   -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
CC       RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
CC       AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
CC       PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
CC       SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE
CC       PROTEASES) MEDIATING APOPTOSIS. CONTRIBUTES TO THE INDUCTION OF
CC       NONCYTOCIDAL TNF EFFECTS INCLUDING ANTI-VIRAL STATE AND ACTIVATION
CC       OF THE ACID SPHINGOMYELINASE.
CC   -!- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC       HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
CC       PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
CC       WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
CC       PROTEINS SUCH AS TRAFS, RIP AND POSSIBLY FADD, ARE RECRUITED TO
CC       TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
CC       ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC       NF-KAPPA B SIGNALING.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- DOMAIN: THE DOMAIN THAT INDUCES A-SMASE IS PROBABLY IDENTICAL TO
CC       THE DEATH DOMAIN. THE N-SMASE ACTIVATION DOMAIN (NSD) IS BOTH
CC       NECESSARY AND SUFFICIENT FOR ACTIVATION OF N-SMASE.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD120a entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd120a.htm".
```
```
DR   EMBL; X55313; CAA39021.1; -.
DR   EMBL; M33294; AAA03210.1; -.
DR   EMBL; M58286; AAA36753.1; -.
DR   EMBL; M75865; AAA61201.1; JOINED.
DR   EMBL; M75866; AAA61201.1; -.
DR   EMBL; M75864; AAA61201.1; JOINED.
DR   EMBL; M60275; AAA36755.1; -.
DR   EMBL; A21522; CAA01558.1; -.
DR   EMBL; A34899; GQHTIU.
DR   EMBL; M63121; AAA36754.1; -.
DR   PIR; A35010; A35010.
DR   PIR; S12057; S12057.
DR   PIR; A38208; A38208.
DR   PDB; 1NCF; 31-JUL-94.
DR   PDB; 1TNR; 31-JUL-94.
DR   PDB; 1EXT; 11-JAN-97.
DR   PDB; 1ICH; 07-DEC-95.
DR   MIM; 191190; -.
DR   PROSITE; PS00652; TNFR_NGFR_1; 3.
DR   PROSITE; PS50050; TNFR_NGFR_2; 3.
DR   PROSITE; PS00017; DEATH_DOMAIN; 1.
```

DR   PFAM; PF00020; TNFR_c6; 4.
DR   PFAM; PF00531; death; 1.
KW   Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Apoptosis;
KW   3D-structure.
FT   SIGNAL   1   21
FT   CHAIN   22   455   TUMOR NECROSIS FACTOR RECEPTOR 1.
FT   CHAIN   41   291   TUMOR NECROSIS FACTOR BINDING PROTEIN 1.
FT   DOMAIN   22   211   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   212   234   POTENTIAL.
FT   DOMAIN   235   455   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   43   196   4 X TNFR-CYS (POTENTIAL).
FT   REPEAT   43   82   TNFR-CYS 1.
FT   REPEAT   83   125   TNFR-CYS 2.
FT   REPEAT   126   166   TNFR-CYS 3.
FT   REPEAT   167   196   TNFR-CYS 4.
FT   DOMAIN   338   348   N-SMASE ACTIVATION DOMAIN (NSD).
FT   DOMAIN   356   441   DEATH DOMAIN.
FT   DISULFID   44   58
FT   DISULFID   59   72
FT   DISULFID   62   81
FT   DISULFID   84   99
FT   DISULFID   102   117
FT   DISULFID   105   125
FT   DISULFID   127   143
FT   DISULFID   146   158
FT   DISULFID   149   166
FT   DISULFID   168   179
FT   DISULFID   182   191
FT   DISULFID   185   195
FT   CARBOHYD   54   54   POTENTIAL.
FT   CARBOHYD   145   145   POTENTIAL.
FT   CARBOHYD   151   151   POTENTIAL.
FT   CARBOHYD   151   151   POTENTIAL.
FT   CONFLICT   412   412   MISSING (IN REF. 4).
FT   CONFLICT   443   446   GPAA -> APP (IN REF. 4).
SQ   SEQUENCE   455 AA;   50494 MW; CE0EA06F CRC32;

Query Match
Best Local Similarity   25.28%;   Score 143; DB 1; Length 455;
Matches   32; Conservative 30; Mismatches 57; Indels 8; Gaps 8;

Qy   1 MALKVLLEQEKTFF-TLLY-LLGYLSCKYTCEESGDCROQEFRDSGNCVPCNQCGPGMEL 58
            :| |||| :| ::| :|: :| :| ::: : :| ||:||| :|
Db   11 LPL-VLLELVGIYPSGVIGLVPHLGDREKRDSV-CPQGKYIHPQNNSICCTKCHKGTYL 68

Qy   59 SKEGFGYGEDAQCVACRLHRFKE-DWGFQCKCPCLDCA-VVNRFQKANCSATSDAICGD 116
            :| |||::|| |: |: | | :| ::| ::| :|:
Db   69 YNDCP-GPGQDTDCRECESGSFTASENHLRHCLSCSKCKREMGQVEISSCTVDRDTVCG- 126

Qy 117 CLPGFYR 123
            | ||
Db 127 CRKNQYR 133

RESULT 8
ID   TNRC_HUMAN   STANDARD;   PRT;   435 AA.
AC   P36941;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR
DE   2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
GN   LTBR OR TNFCR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93252381.
RA   BAENS M., CHAFFANET M., CASSIMAN J.J., DEN BERGHE H., MARYNEN P.;
RT   "Construction and evaluation of a hncDNA library of human 12p
RT   transcribed sequences derived from a somatic cell hybrid.";
RL   Genomics 16:214-218(1993).

RN   [2]
RP   FUNCTION.
RX   MEDLINE; 94225209.
RA   CROWE P.D., VAN ARSDALE T.L., WALTER B.N., WARE C.F., HESSION C.,
RA   EHRENFELS B., BROWNING J.L., DIN W.S., GOODWIN R.G., SMITH C.A.;
RT   "A lymphotoxin-beta-specific receptor.";
RL   Science 264:707-710(1994).
-!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
-!- IMMUNE DEVELOPMENT.
-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
-!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).

DR   EMBL; L04270; AAA36757.1; -.
DR   HSSP; P25942; 1CDF.
DR   MIM; 600979; -.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PROSITE; PS00050; TNFR_NGFR_2; 3.
DR   PFAM; PF00020; TNFR_c6; 4.
KW   Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT   SIGNAL   1   30
FT   CHAIN   31   435   LYMPHOTOXIN-BETA RECEPTOR.
FT   DOMAIN   31   227   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   228   248   POTENTIAL.
FT   DOMAIN   249   435   CYTOPLASMIC (POTENTIAL).
FT   DOMAIN   42   211   4 X TNFR-CYS.
FT   REPEAT   42   81   TNFR-CYS 1.
FT   REPEAT   82   124   TNFR-CYS 2.
FT   REPEAT   125   168   TNFR-CYS 3.
FT   REPEAT   169   211   TNFR-CYS 4.
FT   DISULFID   43   58   BY SIMILARITY.
FT   DISULFID   59   72   BY SIMILARITY.
FT   DISULFID   62   80   BY SIMILARITY.
FT   DISULFID   83   98   BY SIMILARITY.
FT   DISULFID   101   116   BY SIMILARITY.
FT   DISULFID   104   124   BY SIMILARITY.
FT   DISULFID   126   132   BY SIMILARITY.
FT   DISULFID   139   148   BY SIMILARITY.
FT   DISULFID   142   167   BY SIMILARITY.
FT   DISULFID   170   185   BY SIMILARITY.
FT   CARBOHYD   40   40   POTENTIAL.
FT   CARBOHYD   177   177   POTENTIAL.
SQ   SEQUENCE   435 AA;   46709 MW; 203B82DD CRC32;

Query Match
Best Local Similarity   26.38%;   Score 136; DB 1; Length 435;
Matches   26; Conservative 24; Mismatches 41; Indels 8; Gaps 8;

Qy   28 TCESGDCR-QQ-EFRDSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKEDWG 85
            |: |: ||:|| |: |::| :|:   :|: :| | |: : | ::| |::|
Db   37 ASENQTCRDQEKEYE-PQHRICCSRCPPGTYVSAKCS-RI-RDIVCATCAENSYNEHWN 93

Qy   86 FQK-CKPCLDCAVVNRFQK-ANCSATSDAICGDCLPGFY 122
            |: | || | :   ::| || : : ||:|| |
Db   94 YLTICQLCRPCDPVMGLEEIAPCTSKRKTQC-RCQPGMF 131

RESULT 9
ID   FASA_HUMAN   STANDARD;   PRT;   335 AA.
AC   P25445;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
DE   (APO-1 ANTIGEN) (CD95 ANTIGEN).
GN   APT1 OR FAS.

Left entry (Human Fas / APO-1)

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91309137.
RA   ITOH N., YONEHARA S., ISHII A., YONEHARA M., MIZUSHIMA S.I.,
RA   SAMESHIMA M., HASE A., SETO Y., NAGATA S.;
RT   "The polypeptide encoded by the cDNA for human cell surface antigen
RT   Fas can mediate apoptosis.";
RL   Cell 66:233-243(1991).
RN   [2]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 226-240; 269-291 AND 321-335.
RX   MEDLINE; 92268122.
RA   OEHM A., BEHRMANN I., FALK W., PAWLITA M., MAIER G., KLAS C.,
RA   LI-WEBER M., RICHARDS S., DHEIN J., TRAUTH B.C., PONSTINGL H.,
RA   KRAMMER P.H.;
RT   "Purification and molecular cloning of the APO-1 cell surface
RT   antigen, a member of the tumor necrosis factor/nerve growth factor
RT   receptor superfamily. Sequence identity with the Fas antigen.";
RL   J. Biol. Chem. 267:10709-10715(1992).
RN   [3]
RP   STRUCTURE BY NMR OF 218-335.
RX   MEDLINE; 97122332.
RA   HUANG B., EBERSTADT M., OLEJNICZAK E.T., MEADOWS R.P., FESIK S.W.;
RT   "NMR structure and mutagenesis of the Fas (APO-1/CD95) death domain.";
RL   Nature 384:638-641(1996).
CC   -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
CC       ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
CC       RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
CC       SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
CC       ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
CC       CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
CC       APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
CC       INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED
CC       SUICIDE OF MATURE T-CELLS, OR BOTH.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
CC       AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD95 entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd95.htm".
```

```
DR   EMBL; M67454; AAA63174.1; -.
DR   EMBL; X63717; CAA45250.1; -.
DR   PIR; A40036; A40036.
DR   PIR; S24543; S24543.
DR   PDB; 1DDF; 12-NOV-97.
DR   MIM; 134637; -.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PROSITE; PS50050; TNFR_NGFR_2; 2.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PFAM; PF00020; TNFR_c6; 2.
DR   PFAM; PF00531; death; 1.
KW   Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal;
KW   3D-structure.
FT   SIGNAL        1     16
FT   CHAIN        17    335       FASL RECEPTOR.
FT   DOMAIN       17    173       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    174    190       POTENTIAL.
FT   DOMAIN      191    335       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       47    166       3 X TNFR-CYS.
FT   REPEAT       47     83       TNFR-CYS 1.
FT   REPEAT       84    127       TNFR-CYS 2.
FT   REPEAT      128    166       TNFR-CYS 3.
FT   DOMAIN      230    314       DEATH DOMAIN.
FT   CARBOHYD    118    118       POTENTIAL.
FT   CARBOHYD    136    136       POTENTIAL.
SQ   SEQUENCE    335 AA;   37732 MW;  3BFBF973 CRC32;

Query Match
Best Local Similarity   4.8%;   Pred. No. 2.94e-07;
Matches  21;  Conservative  21;  Mismatches 30;  Indels  6;  Gaps  6;

Score 133;  DB 1;  Length 335;

Db   59 CHKPCPPGERKARDCTVN-GDEPDCVPCQEGKEYTDKAHFSSKCRRCRLCDEGHGHLEVEI 117
Qy   49 CNQ-CGPGMELLSKECGFGYGEDAQCVACRLHR-FKEDWGFQ-KCKPCLDCAVVNRFQ-KA 104
Db  118 NCTRTQNTKC-RCKPNFF 134
Qy  105 NCSATSDAICGDCLPGFY 122
```

Right entry

```
RESULT 10

ID   VT2_MYXVL      STANDARD;      PRT;      326 AA.
AC   P29825;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
GN   T2.
OS   Myxoma virus (strain Lausanne).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Leporipoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91335768.
RA   UPTON C., MACEN J.L., SCHREIBER M., MCFADDEN G.;
RT   "Myxoma virus expresses a secreted protein with homology to the tumor
RT   necrosis factor receptor gene family that contributes to viral
RT   virulence.";
RL   Virology 184:370-382(1991).
CC   -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
CC       REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
CC       ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
```

```
DR   EMBL; M95181; AAA46632.1; -.
DR   EMBL; A23729; CAA01688.1; -.
DR   PIR; A40566; GCVZML.
DR   HSSP; P19438; 1TNR.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PROSITE; PS50050; TNFR_NGFR_2; 2.
DR   PFAM; PF00020; TNFR_c6; 2.
KW   Receptor; Glycoprotein; Repeat; Signal.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17    326       TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR.
FT   DOMAIN       27    186       4 X TNFR-CYS.
FT   REPEAT       27     62       TNFR-CYS 1.
FT   REPEAT       63    104       TNFR-CYS 2.
FT   REPEAT      105    147       TNFR-CYS 3.
FT   REPEAT      148    186       TNFR-CYS 4.
FT   CARBOHYD     66     66       POTENTIAL.
FT   CARBOHYD    181    181       POTENTIAL.
FT   CARBOHYD    205    205       POTENTIAL.
FT   CARBOHYD    238    238       POTENTIAL.
SQ   SEQUENCE    326 AA;   35208 MW;  2F059A61 CRC32;

Query Match
Best Local Similarity   29.5%;
Score 130;  DB 1;  Length 326;
```

Matches 28; Conservative 16; Mismatches 42; Indels 9; Gaps 8;

```
Db  24 DRGKCRGNDY-EKDGLC--CTSCPPGSYASRLCG-P-GSDTVCSPCKNETFTASTNHAPA  78
         |  |: ||   | | :: ||  | :|::  :||: |  :|| :|:|| |||  |: |
Qy  30 ESGDCRQOEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACLRHRFKEDWGF-OK  88

Db  79 LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
       LTBR OR TNFCR.
Qy  89 CKPCLD-CAVVNRFQKANCSAISDAICGDCLPGFY 122
```

RESULT 11

```
ID   TNRC_MOUSE     STANDARD;      PRT;   415 AA.
AC   P50284;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
GN   LTBR OR TNFCR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CVB; TISSUE=LUNG;
RX   MEDLINE; 96072804.
RA   FORCE W.R., WALTER B.N., HESSION C., TIZARD R., KOZAK C.A.,
RA   BROWNING J.L., WARE C.F.;
RT   "Mouse lymphotoxin-beta receptor. Molecular genetics, ligand binding,
RT   and expression.";
RL   J. Immunol. 155:5280-5288(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96163885.
RA   NAKAMURA T., TASHIRO K., NAZAREA M., NAKANO T., SASAYAMA S.,
RA   HONJO T.;
RT   "The murine lymphotoxin-beta receptor cDNA: isolation by the signal
RT   sequence trap and chromosomal mapping.";
RL   Genomics 30:312-319(1995).
CC   -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
CC       IMMUNE DEVELOPMENT.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
```

```
DR   EMBL; U29173; AAA68964.1; -.
DR   EMBL; L38423; AAB00846.1; -.
DR   EMBL; U30798; AAA81334.1; -.
DR   HSSP; P25942; 1CDF.
DR   MGD; MGI:104875; Ltbr.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PROSITE; PS50050; TNFR_NGFR_2; 3.
DR   PFAM; PF00020; TNFR_c6; 3.
KW   Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    415       LYMPHOTOXIN-BETA RECEPTOR.
FT   DOMAIN       31    223       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    224    244       POTENTIAL.
FT   DOMAIN      245    415       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       42     81       4 X TNFR-CYS.
FT   REPEAT       42     82       TNFR-CYS 1.
FT   REPEAT       82    124       TNFR-CYS 2.
FT   REPEAT      125    170       TNFR-CYS 3.
FT   REPEAT      171    213       TNFR-CYS 4.
FT   DISULFID     43     58       BY SIMILARITY.
FT   DISULFID     59     72       BY SIMILARITY.
FT   DISULFID     62     80       BY SIMILARITY.
FT   DISULFID     83     98       BY SIMILARITY.
FT   DISULFID    101    116       BY SIMILARITY.
FT   DISULFID    104    124       BY SIMILARITY.
FT   DISULFID    126    132       BY SIMILARITY.
FT   DISULFID    139    150       BY SIMILARITY.
FT   DISULFID    142    169       BY SIMILARITY.
FT   DISULFID    172    187       BY SIMILARITY.
FT   CARBOHYD     40     40       POTENTIAL.
FT   CARBOHYD    179    179       POTENTIAL.
SQ   SEQUENCE   415 AA;  44956 MW;  3C5DD121 CRC32;
```

RESULT 12

Query Match 4.2%; Best Local Similarity 30.3%; Score 130; DB 1; Length 415;

Matches 23; Conservative 16; Mismatches 32; Indels 5; Gaps 5;

```
Db  57 VCCSRCPGEFVFAVCS-R-SQDTVCKTCPHNSYNEHHNHLSTCQLCRPCDIVLGFEEVA 114
        | ::|| ||  :| |: | |:| | |  |  | :  | :|  |: || |||||||||  |
Qy  47 VPCNQCGPGMELSKECGFGYGEDAQCVACLRHRFKEDWG-FQKCKPCLDCAVVNRFQK-A 104

Db 115 PCISDRKAEC-RCQPG 129
        || |:  | ||  ||
Qy 105 NCSAISDAICGDCLPG 120
```

```
ID   TNR1_RAT       STANDARD;      PRT;   461 AA.
AC   P22934;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-R1) (P55).
GN   TNFRSF1A OR TNFR1 OR TNFR-1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91090841.
RA   HIMMLER A., MAURER-FOGY I., KROENKE M., SCHEURICH P., PFIZENMAIER K.,
RA   LANTZ M., OLSSON I., HAUPTMANN R., STRATOWA C., ADOLF G.R.;
RT   "Molecular cloning and expression of human and rat tumor necrosis
RT   factor receptor: chain (p60) and its soluble derivative, tumor
RT   necrosis factor-binding protein.";
RL   DNA Cell Biol. 9:705-715(1990).
CC   -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
CC       RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
CC       AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
CC       PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
CC       SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE
CC       PROTEASES) MEDIATING APOPTOSIS (BY SIMILARITY).
CC   -!- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC       HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
CC       PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
CC       WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
CC       PROTEINS SUCH AS TRAFS, RIP AND POSSIBLY FADD, ARE RECRUITED TO
CC       TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
CC       ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC       NF-KAPPA B SIGNALING (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
```

```
DR   EMBL; M63122; AAA42256.1; -.
DR   PIR; B36555; B36555.
DR   HSSP; P19438; LTNR.
```

```
DR   PROSITE; PS00652; TNFR_NGFR_1; 3.
DR   PROSITE; PS00050; TNFR_NGFR_2; 3.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PFAM; PF00020; TNFR_c6; 4.
DR   PFAM; PF00531; death; 1.
KW   Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Apoptosis.
FT   SIGNAL        1     21
FT   CHAIN        22    461       TUMOR NECROSIS FACTOR RECEPTOR 1.
FT   DOMAIN       22    211       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    212    234       POTENTIAL.
FT   DOMAIN      235    461       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       43    196       4 X TNFR-CYS.
FT   REPEAT       43     82       TNFR-CYS 1.
FT   REPEAT       83    125       TNFR-CYS 2.
FT   REPEAT      126    166       TNFR-CYS 3.
FT   REPEAT      167    196       TNFR-CYS 4.
FT   DOMAIN      344    354       N-SMASE ACTIVATION DOMAIN (NSD).
FT   DOMAIN      363    448       DEATH DOMAIN.
FT   DISULFID     44     58       BY SIMILARITY.
FT   DISULFID     59     72       BY SIMILARITY.
FT   DISULFID     62     81       BY SIMILARITY.
FT   DISULFID     84     99       BY SIMILARITY.
FT   DISULFID    102    117       BY SIMILARITY.
FT   DISULFID    105    125       BY SIMILARITY.
FT   DISULFID    127    143       BY SIMILARITY.
FT   DISULFID    146    158       BY SIMILARITY.
FT   DISULFID    149    166       BY SIMILARITY.
FT   DISULFID    168    179       BY SIMILARITY.
FT   DISULFID    182    191       BY SIMILARITY.
FT   DISULFID    185    195       BY SIMILARITY.
FT   CARBOHYD     54     54       POTENTIAL.
FT   CARBOHYD    151    151       POTENTIAL.
FT   CARBOHYD    201    201       POTENTIAL.
SQ   SEQUENCE   461 AA;  50969 MW;  82F68B08 CRC32;

Query Match             4.2%;  Score 130;  DB 1;  Length 461;
Best Local Similarity   25.9%;
  Matches   28;  Conservative   16;  Mismatches   59;  Indels    5;  Gaps    5;

Qy    44 CPGGKYAHPKNNSICCTKCHKGTYLVSDCPSP-GQETVCEVCDKGTFTASQNHVRQCLSC 102
           |  :|  |:| |  ::|     :|||::|  : |:: ||| ||:|  :: :  | ||
Db    34 CRQGEFRDRSGNCVPCNQCGPGMELSKECGFGYGEDAQCVACRLHRFKEDWGF-QKCKPC  92

Db   103 KTCRKEMFQVEISPCKADMDTVCG-CKKNQFQRYLSETHFQCVDCSPC 149
          |||    | ::  |  | : |  |  :  || :   | | |: | |
Qy    93 LDCA-VVNRFQKANCSAISDAICGDCLPG-FYRKTKLVGFQDMECVPC 138
```

```
RESULT 13
ID   FASA_MOUSE     STANDARD;      PRI;   327 AA.
AC   P25446;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
DE   (APO-1 ANTIGEN) (CD95).
GN   APT1 OR FAS.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92148151.
RA   WATANABE-FUKUNAGA R., BRANNAN C.I., ITOH N., YONEHARA S.,
RA   COPELAND N.G., JENKINS N.A., NAGATA S.;
RT   "The cDNA structure, expression, and chromosomal assignment of the
RT   mouse Fas antigen";
RL   J. Immunol. 148:1274-1279(1992).
RN   [2]
RP   SEQUENCE OF 1-96 FROM N.A.
RX   MEDLINE; 93189576.
RA   ADACHI M., WATANABE-FUKUNAGA R., NAGATA S.;
RT   "Aberrant transcription caused by the insertion of an early
RT   transposable element in an intron of the Fas antigen gene of lpr
RT   mice.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:1756-1760(1993).
RN   [3]
RP   VARIANT LPR.
RX   MEDLINE; 92195401.
RA   WATANABE-FUKUNAGA R., BRANNAN C.I., COPELAND N.G., JENKINS N.A.,
RA   NAGATA S.;
RT   "Lymphoproliferation disorder in mice explained by defects in Fas
RT   antigen that mediates apoptosis.";
RL   Nature 356:314-317(1992).
CC   -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
CC       ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
CC       RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
CC       SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
CC       ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
CC       CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
CC       APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
CC       INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED
CC       SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: DETECTED IN VARIOUS TISSUES INCLUDING THYMUS,
CC       LIVER, LUNG, HEART, AND ADULT OVARY.
CC   -!- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
CC       AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
CC   -!- DISEASE: DEFECTS IN FAS ARE THE CAUSE OF A LYMPHOPROLIFERATION
CC       DISORDER (LPR) RESPONSIBLE FOR LYMPHADENOPATHY AND AUTOANTIBODY
CC       PRODUCTION.
CC   -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M83649; AAA37593.1; -.
DR   EMBL; S56490; AAB25700.1; -.
DR   EMBL; S56485; AAB25700.1; JOINED.
DR   EMBL; S56486; AAB25700.1; JOINED.
DR   PIR; A46484; A46484.
HSSP; P25445; 1DDF.
DR   MGD; MGI:95484; FAS.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PROSITE; PS00050; TNFR_NGFR_2; 2.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PFAM; PF00020; TNFR_c6; 3.
DR   PFAM; PF00531; death; 1.
KW   Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal;
KW   Disease mutation.
FT   SIGNAL        1     21
FT   CHAIN        22    327       FASL RECEPTOR.
FT   DOMAIN       22    169       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    170    186       POTENTIAL.
FT   DOMAIN      187    327       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       43    162       3 X TNFR-CYS.
FT   REPEAT       43     79       TNFR-CYS 1.
FT   REPEAT       80    123       TNFR-CYS 2.
FT   REPEAT      124    162       TNFR-CYS 3.
FT   DOMAIN      222    306       DEATH DOMAIN.
FT   CARBOHYD     43     43       POTENTIAL.
FT   CARBOHYD    114    114       POTENTIAL.
FT   VARIANT     246    246       I -> N (IN LPR).
SQ   SEQUENCE   327 AA;  37418 MW;  22D6DC39 CRC32;

Query Match             4.1%;  Score 127;  DB 1;  Length 327;
Best Local Similarity   26.0%;
  Matches   20;  Conservative   18;  Mismatches   34;  Indels    5;  Gaps    5;

Qy    56 CQPGKYAHPKKVEDCKMN-GGTPTCAPCTEGKEYMDKNHYADKCRRCTLCDEEHGLEVETN 114
           | ::  |  |::   |  ::  |  | ::   ::  :         | ::|         :|
```

```
Qy    49  CNQCGPGMELSKECGPFGYGEDAQCVACRLHR-FKEDWGF-QKCKPCLDCAVVNRFQ-KAN  105
Db   115  CTLTQNTKC-KCKPDFY  130
          |:.| :|::| :| ||
Qy   106  CSATSDAICGDCLPGFY  122

RESULT 14
ID   VC22_VARV    STANDARD;    PRT;    349 AA.
AC   P34015;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   PROTEIN C22/B28 HOMOLOG.
GN   G4R.
OS   Variola virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=INDIA-1967 / ISOLATE IND3;
RX   MEDLINE; 93202281.
RA   SHCHELKUNOV S.N., BLINOV V.M., SANDAKHCHIEV L.S.;
RT   "Genes of variola and vaccinia viruses necessary to overcome the host
RT   protective mechanisms.";
RL   FEBS Lett. 319:80-83(1993).
CC   -!- SIMILARITY: CONTAINS 2 LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X69198; CAA49137.1; -.
DR   EMBL; X67117; CAA47540.1; -.
DR   PIR; D36858; D36858.
DR   PIR; S35987; S35987.
DR   PIR; S46888; S46888.
DR   HSSP; P19438; 1NCF.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PROSITE; PS00050; TNFR_NGFR_2; 2.
DR   PFAM; PF00020; TNFR_c6; 2.
KW   Repeat.
FT   DOMAIN     31    108    2 X TNFR-CYS.
FT   REPEAT     31     66    TNFR-CYS 1.
FT   REPEAT     67    108    TNFR-CYS 2.
SQ   SEQUENCE    349 AA;    38189 MW;    50D0B435 CRC32;

Query Match              4.1%;
Best Local Similarity   28.7%;    Score 128; DB 1; Length 349;
                                  Pred. No. 2.37e-06;
Matches   27;  Conservative   21;  Mismatches   37;  Indels    9;  Gaps    8;

Qy    29  NGKCKDTEYK-RHNLG--CLSCPPGTYASRLCD-S-KTNTQCTPCGSGTFTSRNNHLPAC   83
          :|::||::| :|| :|  :::|: |:  :::|: : ::||::| ::::|:| |  ||:|
Db    31  SGDCRQQEFRDRSGNCVPCNCQGPGMELSKECGFGYGEDAQCVACRLHRFKE-DWGFQKC   89

Qy    84  LSCNGRCNS-NQVETRSCNTIHNRIC-ECSPGYY  115
          ||:|| ||: || ::||: : :|:|| ||||||:
Db    90  KPCLD-CAVVNRFQKANCSATSDAICGDCLPGFY  122

RESULT 15
ID   LMG2_MOUSE    STANDARD;    PRT;    1192 AA.
AC   Q61092;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   LAMININ GAMMA-2 CHAIN PRECURSOR (KALININ/NICEIN/EPILIGRIN 100 KD
DE   SUBUNIT).
GN   LAMC2 OR B2T.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB; TISSUE=LUNG;
RX   MEDLINE; 95188994.
RA   SUGIYAMA S., UTANI A., YAMADA S., KOZAK C.A., YAMADA Y.;
RT   "Cloning and expression of the mouse laminin gamma 2 (B2t) chain, a
RT   subunit of epithelial cell laminin.";
RL   Eur. J. Biochem. 228:120-128(1995).
CC   -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC       IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC       CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC       WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC   -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC       DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC       TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC       COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC       THE GAMMA-2 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/
CC       NICEIN).
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC       MEMBRANES (MAJOR COMPONENT).
CC   -!- TISSUE SPECIFICITY: EPITHELIAL CELLS OF MANY TISSUES, PARTICULARLY
CC       HIGH LEVELS IN TONGUE, HAIR FOLLICLES AND KIDNEY. BASEMENT
CC       MEMBRANES OF THE COLLECTING TUBULES OF KIDNEY AND PANCREAS.
CC   -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC       WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC   -!- DOMAIN: DOMAIN IV IS GLOBULAR.
CC   -!- SIMILARITY: CONTAINS 7.5 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U43327; AAA85256.1; -.
DR   HSSP; P02468; 1TLE.
DR   MGD; MGI:99913; LAMC2.
DR   PROSITE; PS00022; EGF_1; 3.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 5.
DR   PFAM; PF00052; laminin_B; 1.
DR   PFAM; PF00053; laminin_EGF; 5.
KW   Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW   Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT   SIGNAL       1     21    POTENTIAL.
FT   CHAIN       22   1192    LAMININ GAMMA-2 CHAIN.
FT   DOMAIN      28    196    3.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
FT                            III B).
FT   DOMAIN      28     83    LAMININ EGF-LIKE 1.
FT   DOMAIN      84    130    LAMININ EGF-LIKE 2.
FT   DOMAIN     139    186    LAMININ EGF-LIKE 3.
FT   DOMAIN     187    196    LAMININ EGF-LIKE 4 (N-TERMINAL).
FT   DOMAIN     197    382    LAMININ DOMAIN IV.
FT   DOMAIN     383    603    4 X LAMININ EGF-LIKE REPEATS (DOMAIN
FT                            III A).
FT   DOMAIN     383    416    LAMININ EGF-LIKE 4 (C-TERMINAL).
FT   DOMAIN     417    462    LAMININ EGF-LIKE 5.
FT   DOMAIN     463    517    LAMININ EGF-LIKE 6.
FT   DOMAIN     518    573    LAMININ EGF-LIKE 7.
FT   DOMAIN     574    603    LAMININ EGF-LIKE 8 (INCOMPLETE).
FT   DOMAIN     604   1192    DOMAIN II AND I.
FT   DOMAIN     613    711    LAMININ EGF-LIKE 1.
FT   DOMAIN     760    787    COILED COIL (POTENTIAL).
FT   DOMAIN     947    997    COILED COIL (POTENTIAL).
FT   DOMAIN    1140   1179    COILED COIL (POTENTIAL).
FT   SITE       587    589    CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID    84     96    BY SIMILARITY.
```

```
FT   DISULFID      86    102   BY SIMILARITY.
FT   DISULFID     104    113   BY SIMILARITY.
FT   DISULFID     116    128   BY SIMILARITY.
FT   DISULFID     139    150   BY SIMILARITY.
FT   DISULFID     141    155   BY SIMILARITY.
FT   DISULFID     157    166   BY SIMILARITY.
FT   DISULFID     169    184   BY SIMILARITY.
FT   DISULFID     463    471   BY SIMILARITY.
FT   DISULFID     465    482   BY SIMILARITY.
FT   DISULFID     485    494   BY SIMILARITY.
FT   DISULFID     497    515   BY SIMILARITY.
FT   DISULFID     518    532   BY SIMILARITY.
FT   DISULFID     520    539   BY SIMILARITY.
FT   DISULFID     542    551   BY SIMILARITY.
FT   DISULFID     554    571   BY SIMILARITY.
FT   DISULFID     611    611   BY SIMILARITY.
FT   DISULFID     614    614   INTERCHAIN (PROBABLE).
FT   DISULFID    1183   1183   INTERCHAIN (PROBABLE).
FT                             INTERCHAIN (WITH CHAIN BETA-3)
FT                             (PROBABLY).
FT   CARBOHYD     342    342   POTENTIAL.
FT   CARBOHYD     363    363   POTENTIAL.
FT   CARBOHYD     527    527   POTENTIAL.
FT   CARBOHYD     942    942   POTENTIAL.
FT   CARBOHYD    1033   1033   POTENTIAL.
SQ   SEQUENCE    1192 AA;  130287 MW;  76A7C102 CRC32;
```

```
Query Match                           Score 126;  DB 1;  Length 1192;
Best Local Similarity   32.1%;        Pred. No. 5.39e-06;
Matches   35;  Conservative  18;  Mismatches  43;  Indels  13;  Gaps  11;

Db        386  LLGYKGQFCQ-ECASG-YKRDSARLGAFGACVPCN-CQGEGACDPDTGDCYSGDENPDIE  442
               ||||   :     ||  ::::     : ||||||| |-|--|--|----|-:|-----:
Qy         19  LLGYLS--CKVTCESGDCRQQEFR-DRSGNCVPCNQC-GPGM-ELSK-ECGFGYGE-DAQ  71

Db        443  CADCPIGFYNDPHDPRSCKPC-PCH--NGFSCSVMPETEEVVCNNCPPG  488
               | | :    ||||| | | :-:   | -:||  :-|  : :| :|-|   ||
Qy         72  CVACRLHRFKEDWGFQKCKPCLDCAVVNRFQKANCSAISDAICGDCLPG  120
```

Search completed: Tue Apr 18 14:00:34 2000
Job time : 43 secs.